# EL PARAÍSO DE CANTOR

## LA TRADICIÓN CONJUNTISTA EN LA FILOSOFÍA MATEMÁTICA

•

ROBERTO TORRETTI



# EL PARAÍSO DE CANTOR

A Joaquín Cordua y Enrique d'Étigny

SUMARIO

APÉNDICES

# PREFACIO

En los siglos XIX y XX la matemática prolifera y florece como quizás ningún otro quehacer del espíritu. Movidos por la misma riqueza y audacia de sus invenciones, algunos matemáticos notables se ponen a reflexionar sobre la naturaleza y alcance de su actividad. Su reflexión es lo que se llama filosófica, y así la entienden; pero la conducen como matemáticos que son, aunando libertad y rigor, fantasía ubérrima y precisión pedante, en el estilo propio de su disciplina. Esta filosofía matemática de la matemática existe de dos maneras. Por una parte, hay una corriente más o menos unitaria de pensamiento que ejerce una enorme influencia sobre la investigación matemática y ha llegado a dominar la enseñanza universitaria. Esta corriente se autodenomina "clásica", pero la llamaré "conjuntista" porque coloca al centro de la matemática, en una forma u otra, la noción de conjunto y trabaja en fortalecerla. Iniciada por Dedekind (1831-1916) y Cantor (1845-1918), incorpora logros de Frege (1848-1925), Peano (1858-1932), Whitehead (1861-1947) y Russell (1872-1970), y recibe aportes de Hilbert (1862-1943), Zermelo (1871-1953), Tarski (1902-1983), von Neumann (1903-1957), Gödel (1906-1978), Gentzen (1909-1945), y muchos otros. Por otra parte, están los adversarios del conjuntismo —ilustres matemáticos como Kronecker (1823-1891), Poincaré (1854-1912), Brouwer (1881-1955) y Weyl (1885-1955), filósofos como Wittgenstein (1889-1951) y Lorenzen (1915-1994)— que impugnan con poderosas razones sus ideas y prácticas más arraigadas, sin que la masa de los matemáticos les preste mucha atención.

Este libro es una historia razonada de la tradición conjuntista, desde los primeros escritos de Cantor hasta la publicación de los teoremas de Cohen (1963) y su impacto inmediato. Hago una que otra alusión a Kronecker y me ocupo, cuando hace falta, de Poincaré, pero la importante oposición de Brouwer y Weyl sólo la menciono al paso, sin analizarla, porque el propio Hilbert, que explícitamente define su empresa filosófica como una defensa de las matemáticas contra ellos, no se dio el trabajo de estudiarlos. Con todo, para el lector curioso, explico brevemente en el Apéndice XIX algunas ideas

de Brouwer que preceden y motivan la declaración de guerra de Hilbert. En cambio, no me ha parecido oportuno examinar en el presente contexto la fundamentación constructivista del análisis, propuesta inicialmente por Weyl (1918), poco antes de su conversión al brouwerismo, y desarrollada mucho más tarde por Lorenzen (1965) y Bishop (1967).

El libro consta de tres partes, correspondientes a tres etapas en la historia del conjuntismo. El presente volumen contiene las partes 1 y 2. La parte 3 todavía no está escrita. La parte 1, titulada "Conjuntos", se refiere a la fundación de la teoría de conjuntos por Cantor, las paradojas que se le enrostran, y la axiomatización de la teoría por Zermelo (1908) y sus continuadores. La parte 2, titulada "Cálculos" gira en torno al programa de Hilbert para darle a la teoría de conjuntos un fundamento intuitivo incontestable, garantizando así a los matemáticos el disfrute del paraíso que —según frase del mismo Hilbert— Cantor ha creado para ellos. Estudia los antecedentes de dicho programa en las obras de Frege, Peano, Dedekind, Russell y Whitehead, y Skolem; su desarrollo en la década de 1920 por Hilbert y sus seguidores; y el inesperado escollo que le salió al encuentro con los hallazgos de Gödel. La parte 3, "Modelos", examinará la contribución al conjuntismo de los métodos semánticos introducidos desde 1930 por Gödel y Tarski.

Pienso que el libro puede servir como introducción histórica al tema. Para leerlo, no es preciso tener conocimientos previos al respecto, pero sí el hábito de leer definiciones y demostraciones matemáticas. Cualquiera que haya seguido cursos universitarios de matemáticas por más de un año tiene ese hábito en la medida requerida aquí. Por otra parte, creo que una persona acostumbrada a leer prosa filosófica puede adquirirlo directamente en este mismo libro. Supongo, sí, que el lector filósofo que se interese en él habrá hecho estudios de lógica. Por otra parte, confío en que el lector con educación matemática pero sin estudios de lógica podrá extraer del Apéndice IX (pp. 480-502) toda la información requerida.

En general explico cada término técnico la primera vez que lo uso. (El índice analítico permitirá ubicar rápidamente tales explicaciones). Pero la terminología lógica se presenta sistemáticamente en el Apéndice IX, y en el Glosario que sigue a los apéndices defino algunos términos de uso común entre los matemáticos, que los lectores que vienen de la filosofía tal vez desconocen (estos términos se señalan con una †). Con rarísimas excepciones, expresamente señaladas, me atengo a la terminología estándar, aunque

no me parece afortunada en todos los casos. Si las investigaciones lógico-matemáticas a que se refiere este libro hubiesen sido conducidas principalmente por personas de habla castellana, con seguridad estaríamos usando términos más eufónicos o elocuentes. Tal como han sido las cosas, tenemos que arreglárnoslas con palabras tan feas como 'completud' (para nombrar la propiedad de ser o estar *completo*) o tan opacas como 'consistencia' (para decir *ausencia de contradicción*).

Las referencias bibliográficas se dan en forma abreviada. Las abreviaturas se explican en la lista de obras citadas que va al final. Constan, generalmente, del nombre del autor y el año de publicación. Cuando se cita más de una obra de un autor aparecida en el mismo año, la segunda, tercera, etc., se distinguen con las letras 'a', 'b', . . . Cuando el año de la edición citada no informa sobre la cronología de la obra, la referencia abreviada contiene, en vez de la fecha, una sigla alusiva al título.

El libro se publica con el patrocinio de la Universidad Nacional Andrés Bello. Estoy muy agradecido a esta Universidad, y especialmente al Rector Joaquín Barceló, por su decisivo apoyo. Agradezco también a la Editorial Universitaria y en particular al editor, Sr. Braulio Fernández, por su interés en la pronta aparición del libro.

Escribí un borrador, muy próximo ya al presente texto, entre 1990 y 1994, cuando aún ocupaba una cátedra de filosofía en el Recinto de Río Piedras de la Universidad de Puerto Rico. Hago público aquí mi agradecimiento al Rector y la Junta Universitaria por la liberalidad con que me concedieron el tiempo necesario para este proyecto. Doy asimismo las gracias a las bibliotecarias y bibliotecarios del Recinto que adquirieron y catalogaron de urgencia publicaciones nuevas que me hacían falta, hicieron venir en préstamo desde bibliotecas del continente otras ya agotadas, y con invariable amabilidad y diligencia atendieron a mi voraz demanda de fotocopias.

Jorge López Fernández, Francisco Rodríguez Consuegra y Matthias Schirn leyeron el susodicho borrador y propusieron importantes mejoras. Atendiendo a las indicaciones de Rodríguez Consuegra redacté nuevamente buena parte del capítulo sobre la teoría de los tipos de Russell (§ 2.4), y siguiendo el consejo de Schirn amplié bastante —aunque no tanto, quizás, como él quería— la discusión filosófica del programa de Hilbert (§ 2.9). Doy aquí las gracias a estos buenos amigos por su valiosa ayuda. Como ninguno de ellos

ha visto la versión final, no puede caberles responsabilidad alguna por los errores que todavía contiene.

Este libro, al igual que otros que he publicado anteriormente, no habría sido posible sin el apoyo continuo de Carla Cordua. Por la índole del texto, esta vez fueron pocos los pasajes que le infligí a medio redactar, pidiéndole que me puliera el estilo. Pero es claro que, de no ser por ella, todo este trabajo no me hubiera valido la pena.

Aunque mi afición a la filosofía de las matemáticas (y la física) data de los años cuarenta, no habría alcanzado nunca la intensidad que se advertirá en estas páginas, si Enrique d'Étigny y Joaquín Cordua no hubieran tenido la idea de llevarme, en los años sesenta, a enseñar filosofía a la Facultad de Ciencias Físicas y Matemáticas de la Universidad de Chile. En recuerdo de esos buenos tiempos, les dedico el libro, con gratitud y afecto.

*Santiago de Chile, 31 de mayo de 1998.*

# 1

# CONJUNTOS

## 1.1 LA PALABRA 'CONJUNTO' EN LA MATEMÁTICA DEL SIGLO XX

La palabra 'conjunto' (alemán, 'Menge'; francés, 'ensemble'; inglés, 'set') figura destacadamente en la literatura matemática contemporánea. Los cursos y manuales universitarios suelen presentar su respectivo campo como una especie de "conjuntos", caracterizada por ciertas condiciones que han de cumplir los "elementos" —por lo demás indeterminados— de cualquier "conjunto" de esa especie, o de ciertos otros "conjuntos" generados desde él mediante operaciones estándar. El estudio de ese campo consiste entonces en determinar las consecuencias de tales condiciones generales, o de otras, más estrechas, mediante las cuales se caracterizan una o más subespecies interesantes de la especie inicial.[1]

Aunque los matemáticos ceden de buen grado a la tentación de usar palabras corrientes del idioma en acepciones peculiarísimas inventadas por ellos, parece que toman la palabra 'conjunto' en su significado habitual o, en todo caso, en uno muy próximo a éste.[2] Salvo por las dos excepciones que comentaré en los próximos párrafos, la palabra 'conjunto' normalmente designa en la literatura matemática una colección de objetos de cualquier clase —los

[1] Véanse en el Glosario las definiciones de 'espacio topológico', 'grupo', 'grupo abeliano'. En lo sucesivo, el símbolo † a continuación de una palabra, o al comienzo y al final de una expresión, sirve para advertir que esa palabra o expresión está explicada en el Glosario.

[2] En su útil *Introducción filosófica a la teoría de conjuntos* (1990), Stephen Pollard combate este parecer. Según él, "es una locura" suponer que los conjuntos de los matemáticos son objetos familiares a todo el mundo (p. 12) y la creencia de que "el desarrollo de la teoría matemática de conjuntos ha sido significativamente influenciado por nociones tomadas directamente del pensamiento cotidiano" es un "mito" (p. 14). Como reacción contra la manía pedagógica de enseñarle teoría de conjuntos a los párvulos estas aseveraciones de Pollard son muy comprensibles, pero pervierten los hechos. El geómetra que habla del grupo formado por el *conjunto* de las simetrías del cubo y el aficionado a la música que habla de tal o cual *conjunto* instrumental no emplean la palabra 'conjunto' en acepciones radicalmente diferentes.

*elementos* del *conjunto*— reunidos en la realidad o en el pensamiento del estudioso. 'Objeto' debe entenderse aquí en el sentido más amplio posible, sin hacer diferencias entre lo real y lo imaginario, o entre lo sustancial y lo accidental. La reunión de tales o cuales elementos en un mismo conjunto puede fundarse en una propiedad común o en una relación entre ellos, pero no tiene que ser así. El conjunto de las tres virtudes teologales es sin duda más interesante, pero no más legítimo, matemáticamente hablando, que aquél cuyos elementos son el perro de Las Meninas, el Peñón de Gibraltar y la raíz cuadrada de 2. La identidad de un conjunto depende total y exclusivamente de la identidad de sus elementos, la cual tiene que estar cabal y exactamente determinada para que el conjunto esté definido. Esencial es que el conjunto formado reuniendo determinados objetos constituya a su vez un objeto —en el amplio sentido indicado— y pueda, por lo tanto, entrar como elemento en un nuevo conjunto.[3] Los matemáticos distinguen, sí, entre los elementos *contenidos* en un conjunto y las *partes* (*subconjuntos*) *incluidas* en él. Por definición, $x$ es parte de $y$ si y sólo si todo elemento $z$ de $x$ es también un elemento de $y$ (simbólicamente: $x \subseteq y \leftrightarrow (z \in x \rightarrow z \in y)$). Obsérvese que según esta definición todo conjunto es una parte de sí mismo, $y \subseteq y$. Si $x \subseteq y$ pero $x \neq y$ decimos que $x$ es una *parte propia* de $y$ (simbólicamente: $x \subset y$). En la conversación ordinaria no distinguiríamos con tanto celo entre 'contener un elemento' e 'incluir una parte', pero, una vez que se lo explica, el distingo ciertamente parece razonable.

Paso a describir las dos excepciones a que aludí. Considérese un conjunto con, digamos, cuatro elementos, $a$, $b$, $c$ y $d$. Conforme a la práctica usual lo llamaré el conjunto $\{a,b,c,d\}$. Los conjuntos $\{b,c\}$ y $\{b,c,d\}$ son obviamente partes de $\{a,b,c,d\}$. El *complemento* de $\{b,c\}$ en $\{a,b,c,d\}$ es el conjunto $\{a,d\}$ formado al quitar de $\{a,b,c,d\}$ los elementos de $\{b,c\}$ (simbólicamente: $\{a,b,c,d\}\setminus\{b,c\} = \{a,d\}$). ¿Cuál es el complemento de

[3] Como se puede ver, la noción matemática de *conjunto* nada tiene que ver con la noción lógica de *clase* o *extensión de un concepto*. Contrastando esta última con lo que Schröder llamó 'dominio' (*Gebiet*) —y que corresponde casi exactamente a nuestra noción de 'conjunto'— Frege (1895, p. 455) escribe: "Considero fallido el intento de basar la extensión del concepto como clase no sobre el concepto sino sobre los objetos individuales. [ . . . ] La extensión de un concepto no consta de los objetos que caen bajo el concepto, como un bosque de árboles, sino que tiene en el concepto mismo su único sostén (*Halt*). Así, el concepto tiene primacía lógica sobre su extensión."

$\{b,c,d\}$ en $\{a,b,c,d\}$? Para el matemático la respuesta natural es $\{a\}$, el conjunto cuyo único elemento es $a$. Tal respuesta desafía el uso ordinario del castellano, que respetábamos hasta hace un momento, cuando decíamos que un conjunto era una colección de objetos (así, en plural). Por otra parte, hablar de conjuntos de un solo elemento o conjuntos *unitarios* (como diré en lo sucesivo) no es un abuso de lenguaje muy grave y tiene la ventaja de permitirnos definir la complementación (esto es, la operación de tomar el complemento) como una †operación algebraica† sobre el conjunto de las partes de un conjunto dado.

La segunda excepción es más inquietante: la matemática contemporánea acepta unánimemente un objeto al que llama “conjunto” pero que no tiene elementos: el *conjunto vacío*, designado por el símbolo $\varnothing$. Es evidente que, si $\varnothing$ existe, está incluido en todos los conjuntos (puesto que, si $x$ es un conjunto cualquiera, no hay ningún elemento de $\varnothing$ que no sea a la vez un elemento de $x$). Este resultado, en virtud del cual todos los conjuntos, por heterogéneos que sean, incluyen una “parte” común, exacerba la desconfianza que inspira al sano sentido común la idea misma de un conjunto sin elementos. Me parece que el principal motivo que induce a los matemáticos a aceptar el conjunto vacío $\varnothing$ es algebraico: en presencia de $\varnothing$ —y sólo gracias a ella— es posible conferir la estructura de un †álgebra de Boole† al conjunto de las partes de un conjunto cualquiera $K$ (en adelante, $\mathcal{P}K$). Consideremos la operación de intersección que asigna a dos conjuntos cualesquiera el conjunto de los elementos comunes a ambos. Volviendo a nuestro ejemplo del conjunto $\{a,b,c,d\}$, comprobamos que la intersección de sus partes $\{b,c\}$ y $\{a,c,d\}$ es el conjunto unitario $\{c\}$, que también es una parte de $\{a,b,c,d\}$ (simbólicamente: $\{b,c\} \cap \{a,c,d\} = \{c\}$). ¿Cuál es, empero, la intersección de dos partes disjuntas de $\{a,b,c,d\}$ (esto es, de dos partes que no tienen un elemento común)? ¿Diremos que la ecuación $x = \{b,c\} \cap \{a,d\}$ no tiene solución? ¿que la intersección *no está definida* sobre pares disjuntos? Eso es lo que diría, seguramente, cualquier persona sin educación matemática. Pero el matemático puede eludir esta incómoda irregularidad gracias a que acepta la existencia del conjunto sin elementos $\varnothing$, el cual por definición es parte de cualquier conjunto y constituye la intersección de cualquier par de conjuntos disjuntos. La matemática había dado ya pasos aparentemente mucho más audaces cuando aceptó la existencia de soluciones para ciertas ecuaciones numéricas tales como $x = 5 - 9$, $8x = 3$, $x^2 = 2$, ó $x^2 + 1 = 0$.

Por otra parte, no se trata de conjurar la existencia de un objeto inexistente por un acto de hechicería intelectual.[4] Basta fijar cualquier objeto que no sea un conjunto y que por lo tanto no tenga elementos —por ejemplo, el Polo Sur o la ira de Aquiles— y decidir que se lo llamará 'conjunto'. Con esta sola sencillísima convención hacemos de ese objeto una parte (vacía) de todo conjunto, que funciona como cero en el álgebra de Boole definible en el conjunto de todas las partes de un conjunto dado: Si $k$ es una de esas partes, tenemos que $k \cup \varnothing = k$ y $k \cap \varnothing = \varnothing$.[5]

Históricamente hubo también otro motivo —aunque inválido— para aceptar el conjunto vacío $\varnothing$. Para entender y apreciar este motivo hay que tener presente que un conjunto puede identificarse dando una lista de sus elementos (como hicimos con $\{a,b,c,d\}$) o especificando una o más condiciones que satisfacen todos los miembros del conjunto y sólo ellos. Simbólicamente, el conjunto de todos los objetos $x$ que cumplen una condición $K$ se escribe '$\{x: Kx\}$'. Por ejemplo, $\{x{:}x$ es un número primo y $x < 16\} = \{2,3,5,7,11,13\}$. En las postrimerías del siglo XIX algunos filósofos pensaron por eso que un "conjunto", en el sentido en que esta palabra empezaba a usarse en ciertas publicaciones matemáticas, era lisa y llanamente lo mismo que en lógica se conoce como la *extensión* de un concepto. (Evidentemente, los números contenidos en $\{2,3,5,7,11,13\}$ constituyen la extensión del concepto 'número primo menor que 16'). Visto de este modo, el conjunto $\varnothing$ coincide con la

4 Aunque algunos grandes matemáticos se han expresado como si se tratara justamente de eso. Así Dedekind (1888, p. 2) anuncia que en ese escrito no admitirá el conjunto vacío "aunque para otras investigaciones puede ser cómodo inventarlo". ("[Wir wollen] das leere System, welches gar kein Element enthält, aus gewissen Gründen hier ganz ausschließen, obwohl es für andere Untersuchungen bequem sein kann, ein solches zu erdichten".) En la primera axiomatización de la teoría de los conjuntos, Ernst Zermelo se toma esta libertad. Postula allí el siguiente:

> **Axioma II.** Hay un conjunto (impropio), el "conjunto cero" 0, que no contiene ningún elemento.
>
> (Zermelo 1908a, p. 263)

Lewis (1991) arbitrariamente *define* el conjunto vacío $\varnothing$ como la fusión de todos los objetos individuales existentes (esto es, el objeto individual que comprende a todos los demás dentro de sí). Esta definición asegura que $\varnothing$ existe y está bien determinado como quiera que esté constituido el universo.

5 George Boole introdujo el álgebra que lleva su nombre como una estructura discernible en el sistema lógico de las clases (extensiones de conceptos). Para ello tuvo

extensión de cualquier concepto contradictorio ($\varnothing = \{x{:}x \neq x\} = \{x{:}x$ es un cuadrado redondo$\}$, etc.). Pero, como veremos en el Capítulo 1.6, si bien todo conjunto identificado mediante una condición tiene como elementos exactamente los objetos que caen en la extensión del concepto definido por ella, no toda condición expresable en buen castellano identifica un conjunto. No puede, entonces, darse por descontado que haya un conjunto que identifican las condiciones contradictorias.

---

que completar dicho sistema, por convención terminológica, con dos “objetos” que normalmente no se habrían considerado como clases, a saber, el “universo” y la “clase sin miembros”:

> Por clase se entiende usualmente una colección de individuos, a cada uno de los cuales se puede aplicar un nombre o descripción particular; pero en esta obra el significado del término se extenderá de modo que incluya el caso en que no existe más que un solo individuo que responde al nombre o descripción requeridos, así como los casos denotados por los términos “nada” y “universo”, los cuales, considerados como “clases”, debe entenderse que comprenden, respectivamente, “ningún ente” (*“no beings”*) y “todos los entes”.
>
> (Boole 1854, p. 28)

## 1.2 ‘CONJUNTO’ (‘MENGE’) EN EL VOCABULARIO DE CANTOR

Desde una perspectiva actual, podemos ver el conjuntismo en acción ya desde el comienzo mismo de la matemática moderna en la *Geometría* de Descartes (1637). Al caracterizar cada figura geométrica por la ecuación que satisfacen las coordenadas de sus puntos, la geometría cartesiana representa en efecto la figura mediante el conjunto de los puntos incidentes en ella, seleccionados de entre todos los puntos del espacio por la condición impuesta a sus coordenadas. Pero sólo a fines del siglo XIX y principios del XX el enfoque conjuntista se hará sentir con toda su fuerza gracias a la obra de Georg Cantor y a la influencia que ejerce sobre las siguientes generaciones de matemáticos.

En el Capítulo 1.3 veremos cómo un problema clásico de la matemática decimonónica llevó a Cantor a considerar ciertos conjuntos de puntos de la recta, y de ahí a la investigación de tales conjuntos en general. Les dedica seis trabajos “Über unendliche lineare Punktmannigfaltigkeiten” (“Sobre variedades lineales infinitas de puntos”) publicados en *Mathematische Annalen* entre 1879 y 1884. Ya en el tercero de ellos introduce consideraciones sobre “una variedad (un agregado, un conjunto) de elementos pertenecientes a cualquier esfera conceptual” (Cantor, GA, p. 150; citado *infra*), y el quinto versa sobre los “Fundamentos de una teoría general de las variedades”, título bajo el cual circula como folleto separado. Como puede verse por la frase recién citada, Cantor usa la palabra ‘conjunto’ (*Menge*) como un sinónimo de lo que solía llamarse en su tiempo ‘variedad’ o ‘multiplicidad’ (*Mannigfaltigkeit*).[1] Otros sinónimos utilizados por él son *Gesamtheit* (‘totalidad’) e *Inbegriff* (arriba traducido ‘agregado’). La connotación de estos vocablos se

[1] La matemática alemana toma el término ‘Mannigfaltigkeit’ del vocabulario filosófico de Kant, probablemente a través de Fries. Kant llamaba así a la variedad concreta de los datos de los sentidos, y también a la variedad, pluralidad o multiplicidad comprendidas en el espacio de la geometría, por una parte, y en el tiempo de la mecánica, por otra. En su lección inaugural *Sobre las hipótesis que están en la base de la geometría* (1854), Riemann usa el término con la misma amplitud con que luego lo empleará Cantor. Riemann distingue entre “variedades discretas” (*diskrete Mannigfaltigkeiten*),

explica al comienzo del último trabajo que Cantor dedica al tema, "Contribuciones a la fundamentación de la teoría de los conjuntos transfinitos" (Cantor 1895/97):

> Unter einer "Menge" verstehen wir jede Zusammenfassung *M* von bestimmten wohlunterschiedenen Objekten *m* unsrer Anschauung oder unseres Denkens (welche die "Elemente" von *M* genannt werden) zu einem Ganzen.
>
> Entendemos por 'conjunto' cualquier reunión en un todo *M* de determinados objetos bien distinguidos *m* de nuestra intuición o nuestro pensamiento (llamados 'elementos' de *M*).
>
> (Cantor, GA, p. 282)[2]

---

que constan de "elementos", y "variedades continuas" (*stetige Mannigfaltigkeiten*), que constan de "puntos". Sin embargo, en la literatura matemática actual, el término alemán y sus equivalentes en otros idiomas (E. 'variedad'; F. 'variété'; I. 'manifold'; IT. 'varietà') se usan exclusivamente para referirse a objetos afines a las variedades continuas de Riemann.

2 En "Fundamentos de una teoría general de las variedades", § 1, nota 1, Cantor había dicho:

> Por "variedad" (*Mannigfaltigkeit*) o "conjunto" (*Menge*) entiendo en general cualquier pluralidad que se deja concebir como unidad (*jedes Viele, welches sich als Eines denken läßt*), es decir, cualquier agregado (*Inbegriff*) de elementos determinados que en virtud de una ley pueden ser combinados en un todo.
>
> (Cantor, GA, p. 204)

Bernhard Bolzano había utilizado el término 'Menge' en una acepción similar en su obra póstuma *Paradojas del Infinito* (1851):

> Einen Inbegriff, den wir einem [ . . . ] Begriffe unterstellen, bei dem die Anordnung seiner Teile gleichgültig ist [ . . . ], nenne ich eine Menge.
>
> Llamo conjunto a un agregado subordinado a un concepto, cuando no importa el orden de sus partes.
>
> (Bolzano 1964, p. 4)

El *concepto* que gobierna la formación de un conjunto en el sentido de Bolzano puede equipararse a la *ley* que según Cantor preside la combinación de los elementos del conjunto en un todo. Es significativo que en sus caracterizaciones más tardías de la noción de conjunto Cantor no haga alusión a la existencia de una ley tal. Veo aquí una evolución comparable a la del concepto matemático de función de d'Alembert a Dirichlet (esbozada al comienzo del Capítulo 1.3).

La frase “reunión en un todo” expresa, me parece, que un conjunto es él mismo un objeto, “concebido como una cosa de por sí”, como dice Cantor en un pasaje paralelo (Cantor, GA, p. 411). La indicación de que consta de objetos de “nuestra intuición o nuestro pensamiento” no debe interpretarse como una restricción encaminada a subordinar la noción de conjunto al alcance de nuestras facultades mentales. Se dirige más bien a subrayar que la índole variopinta de los objetos reunidos en un conjunto en nada afecta su viabilidad. En el citado pasaje paralelo expresa la misma idea diciendo que dichos objetos pueden ser “cosas concretas o conceptos abstractos”. Estas dos categorías ontológicas presumiblemente agotaban el reino de lo posible a ojos de Cantor. Ahora bien, “intuición” y “pensamiento” son las facultades cognitivas tradicionalmente asociadas por la filosofía alemana a lo concreto y lo abstracto. Por otra parte, es esencial que los elementos de un conjunto estén exactamente determinados y no se confundan entre ellos, ni con otras cosas que no son elementos del conjunto. En el pasaje arriba aludido del tercer trabajo sobre las variedades lineales infinitas de puntos, esto se traducía en la exigencia de que los conjuntos bajo consideración estuviesen “bien definidos” (*wohldefiniert*):

> Llamo *bien definida* una variedad (una totalidad, un conjunto) de elementos pertenecientes a cualquier esfera conceptual si sobre la base de su definición y como consecuencia del principio lógico del tercero excluido hay que considerar *internamente determinado*, por una parte, si un objeto cualquiera de la misma esfera conceptual pertenece o no como elemento a dicha variedad, y, por otra, si dos objetos pertenecientes al conjunto, no obstante diferencias formales en el modo como son dados, son o no iguales entre sí.
>
> (Cantor, GA, p. 150).

Cantor agrega que la decisión acerca de si un objeto dado $a$ pertenece o no a un conjunto bien definido $M$, o es o no idéntico a un objeto dado $b$ “generalmente no puede efectuarse con seguridad y precisión mediante los métodos y aptitudes disponibles”. Pero lo que importa no es esto, sino “sólo la *determinación interna*, que en casos concretos, cuando los fines buscados lo requieran, se articulará, perfeccionando los medios auxiliares, como una *determinación efectiva* (*externa*).” Cantor nada dice sobre los medios auxiliares que permiten articular la determinación efectiva de los elementos de

un conjunto bien definido; pero supongo que en todo caso serán variantes de los dos que mencioné al final del Capítulo 1.1: o bien se da una lista de todos los elementos del conjunto, nombrando o describiendo a cada uno de una manera inequívoca; o bien se establece una condición que cumplen todos los elementos del conjunto y sólo ellos y un procedimiento para decidir si un objeto cualquiera satisface o no dicha condición. El requisito de "determinación interna" podría entonces suponerse cumplido con sólo prescribir tal condición, aunque no se conozca el procedimiento de decisión correspondiente. La caracterización de 'conjunto' que trascribí de las "Contribuciones" (GA, p. 282) significa, a mi modo de ver, que la teoría general de los conjuntos concierne solamente a conjuntos bien definidos, en este sentido.

En suma, un conjunto en el sentido de Cantor es un objeto constituido por otros objetos —los elementos del conjunto— de tal modo que su identidad depende de la determinación precisa de cuáles objetos son elementos suyos y cuáles no.[3] Esta noción de conjunto concuerda en lo esencial con la expuesta en el Capitulo 1.1. De hecho, a pesar de su explícita insistencia en que los conjuntos constan de "elementos bien distinguidos", Cantor se refiere ocasionalmente a un conjunto que consiste de un solo elemento (GA, p. 98). En cambio, su postura con respecto al conjunto vacío es menos clara. En el Nº 1 de la serie "Sobre variedades lineales infinitas de puntos" dice que "conviene tener un símbolo que exprese la ausencia de puntos, para lo cual elegimos la letra *O*. $P = O$ significa entonces que el conjunto de puntos *P* no contiene *ni un solo* punto, o sea que, estrictamente hablando, no existe como tal" (GA, p. 146). Así, pues, al mismo tiempo que niega la existencia del conjunto vacío, Cantor le pone un nombre, lo cual es quizás la vía más

3 La misma noción de conjunto es adoptada por Dedekind en el influyente ensayo que dedica por esos años a la fundamentación conjuntista de la aritmética (aunque Dedekind dice 'sistema' —'System'— en vez de 'conjunto' —'Menge'):

> Ocurre muy a menudo que diversas cosas *a*, *b*, *c* . . . por algún motivo son concebidas bajo un punto de vista común y reunidas en la mente. Se dice entonces que forman un sistema *S*. [ . . . ] Como objeto de nuestro pensamiento, tal sistema *S* (o sea una colección, o una variedad, o una totalidad [*oder ein Inbegriff, oder Mannigfaltigkeit, oder Gesamtheit*]) también es una cosa; está completamente determinado, cuando está determinado respecto de cada cosa si es o no un elemento de *S*.
>
> (Dedekind 1888, pp. 1s.)

segura —en matemáticas como en literatura— para darle realidad a una ficción. En el Nº 6 de la misma serie da un paso más hacia el reconocimiento del conjunto ∅. Cantor considera allí un conjunto de puntos $P$ incluido en una región $H$ de un espacio $n$-dimensional $G$. Si $H$ se descompone en un número finito o infinito de regiones conexas disjuntas $H_1, H_2, \dots H_k, \dots$, $P$ también se divide en “un número correspondiente” de partes disjuntas $P_1$, $P_2, \dots P_k, \dots$, donde $P_\nu = H_\nu \cap P$ ($\nu = 1, 2, \dots$). Cantor comenta que “$P_\nu$ puede, entonces, ser igual a *cero* (*Null*), en caso de que ningún punto de $P$ caiga dentro de la región $H_\nu$” (GA, p. 210). Cantor no dice que el cero que aquí se nombra sea un conjunto, pero lo ve sin duda como un objeto por derecho propio, y además lo equipara a una parte de un conjunto. El lector advertirá que el ∅ se presenta aquí justamente como la intersección de dos conjuntos disjuntos. El pasaje sugiere, además, que la letra $O$ se le impuso a Cantor como símbolo cuando escribía el Nº1 por su parecido con el número 0.

## 1.3 SERIES TRIGONOMÉTRICAS

El problema matemático que lleva a Cantor a ocuparse con determinados conjuntos de puntos de la recta concierne a la representación de funciones reales mediante series trigonométricas. Una función real es una aplicación† cuyos argumentos y valores son números reales. Las definiciones modernas de número real datan de la época de Cantor, y como él mismo fue el autor de una de ellas, es de suponer que no le contentaron las anteriores. Luego bosquejaré la teoría de los números reales de Cantor. Entre tanto, para no pecar de excesivo anacronismo, entenderemos que una función real es una correspondencia $f$ entre las distancias (orientadas) determinables en dos rectas $\lambda_1$ y $\lambda_2$ en que se ha marcado un punto 0 y un punto a distancia +1 del 0. Diremos que tales rectas han sido *parametrizadas*. $f$ asigna a cada distancia $x$ medida desde el 0 en la recta parametrizada $\lambda_1$ una distancia $y = f(x)$ medida desde el 0 en $\lambda_2$. Como hay un y sólo un punto sobre cada recta a una dada distancia (orientada) del respectivo 0, $f$ se concibe naturalmente como una correspondencia entre puntos, una aplicación del conjunto de los puntos situados en $\lambda_1$ en el conjunto de los puntos situados en $\lambda_2$. (La función $f$ puede también estar definida sólo sobre una parte propia —por ejemplo, un intervalo— de la recta $\lambda_1$.)

Aunque los matemáticos anteriores a Cantor no se expresaban en estos términos, es difícil eludir la impresión de que pensaban de esta manera. En el siglo XVIII, algunos autores, como d'Alembert, entienden al parecer que una función real $f$ tiene que definirse mediante operaciones algebraicas sobre la "variable independiente" $x$ (y constantes); mientras que otros, más atentos a las aplicaciones físicas, piensan que ella puede ser cualquier correspondencia $x \mapsto f(x)$ que represente una curva plana arbitraria (conforme a la convención habitual que mide los argumentos $x$ sobre el eje horizontal de las abscisas y los valores $y = f(x)$ sobre el eje vertical de las ordenadas). Estas dos nociones —algebraica y geométrica— de una función real son demasiado estrechas en comparación con la idea generalizada de función que Dirichlet introduce en 1837:

> Sean $a$ y $b$ dos valores fijos y sea $x$ una cantidad variable tal que $x$ toma paulatinamente todos los valores entre $a$ y $b$. Si un único $y$ finito corresponde a cada $x$ de modo que, mientras $x$ recorre continuamente el intervalo desde $a$ hasta $b$, $y = f(x)$ también varía gradualmente, se dice que $y$ es una función continua o regular de $x$ sobre este intervalo. No es necesario que $y$ esté sujeta a la misma regla con respecto a $x$ a través de todo el intervalo. Ni siquiera es necesario que la relación pueda expresarse mediante operaciones matemáticas.[1]

Dirichlet admite además que la variación paulatina concomitante de $x$ y $f(x)$ se interrumpa para uno o más valores de $x$. En otras palabras, $f$ puede tener una o más discontinuidades. La matemática posterior fue aún más liberal: la continuidad, interrumpida o no, no es un rasgo esencial del concepto de función.

Es de suponer que Joseph Fourier partía de la idea geométrica de función cuando anunció en 1807 que cualquier función $f$ definida en un intervalo finito de la recta real —digamos, en $[-\pi, \pi]$— puede representarse mediante una serie trigonométrica de este modo:

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty}\left(a_n \cos nx + b_n \operatorname{sen} nx\right) \tag{1}$$

donde los coeficientes $a_n$ (para $n \geq 0$) y $b_n$ (para $n \geq 1$) están dados por:

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos nx\,\mathrm{d}x \tag{2}$$

y

$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\operatorname{sen} nx\,\mathrm{d}x \tag{3}$$

Fourier demostró su tesis para algunos casos especiales y esbozó, con escaso rigor, una demostración general.[2] Años más tarde, Dirichlet, en el escrito

[1] Dirichlet, "Über die Darstellung ganz willkürlicher Functionen durch Sinus und Cosinus-reihen", *Repertorium der Physik*, Berlin: Veit, 1837, vol. I, p. 152; citado por Mannheim 1964, pp. 52s.

[2] Fourier, *Théorie analytique de la chaleur* (1822).

arriba citado, estableció ciertas condiciones suficientes para que una función $f$ pueda representarse de este modo y Riemann, en su *Habilitationschrift* de 1854, intentó establecer condiciones necesarias. En este trabajo, Riemann dio la definición de integral que todavía sirve de base a la enseñanza elemental del cálculo y demostró varios resultados importantes. Pero la cuestión de las condiciones necesarias y suficientes de la representación (1) quedó —y sigue— pendiente.

Aunque no se sepan exactamente las condiciones en que una función real admite la representación (1), es importante saber si esa representación es única, cuando existe, o si, por el contrario, una misma función puede representarse mediante dos o más series trigonómetricas diferentes. Esta es la cuestión que Cantor abordó y resolvió en varios artículos publicados entre 1870 y 1872.[3] Cantor (1870) considera una función real $f$ definida en un intervalo finito **I** y representable mediante una serie trigonométrica de la forma (1), convergente para todo $x \in \mathbf{I}$. No presupone que $f$ sea integrable o que los coeficientes de la serie (1) estén dados por (2) y (3). Tampoco requiere la convergencia uniforme de la serie (1) en **I**.[4] La unicidad de la representación (1) es un corolario del siguiente teorema:

[T1] La serie trigonométrica en el lado derecho de la ecuación (1) representa la función constante $f(x) = 0$ si y sólo si $a_n = b_n = 0$ para todo número no negativo $n$.

En efecto, si hay dos representaciones de una misma función $g$ mediante series de la forma (1), la diferencia de las dos series, tomada término a término, representa la función $f(x) = 0$. Pero esto supone, según el teorema citado, que la diferencia entre coeficientes homólogos de ambas series sea en

3 Los trabajos de Cantor aludidos a continuación —como todos sus escritos más importantes— han sido resumidos en inglés por Dauben (1979), pero en esta parte, como en otras, la exposición de Dauben tiene mucho que envidiarle a la claridad del original. J. Marshall Ash (1989) ofrece una demostración rigurosa y detallada de los resultados de Cantor (1870).

4 Decimos que la serie $\sum_{k=1}^{\infty} f_k(x)$ *converge uniformemente* al límite $f(x)$ en el intervalo **I**, si para cada número real positivo ε hay un entero positivo $N(\varepsilon)$ tal que, para todo $x \in \mathbf{I}$ y todo $n > N(\varepsilon)$, $|f(x) - \sum_{k=1}^{n} f_k(x)| < \varepsilon$.

cada caso igual a 0 o, en otras palabras, que las dos representaciones sean idénticas. El teorema T1 se deduce fácilmente del siguiente resultado, que Cantor demuestra utilizando un ingenioso artificio aprendido de Riemann:

[T2] Si para todo $x$ en el intervalo finito **I**,

$$\lim_{n\to\infty}(a_n \operatorname{sen} nx + b_n \cos nx) = 0 \tag{4}$$

$$\text{entonces, } \lim_{n\to\infty} a_n = 0 \quad \text{y} \quad \lim_{n\to\infty} b_n = 0\,.$$

En trabajos posteriores Cantor simplifica la demostración de esta proposición conforme a una sugerencia que le hizo Kronecker y establece que ella sigue siendo válida aunque haya valores excepcionales de $x$ para los cuales la secuencia (4) no converge en absoluto o converge a un límite distinto de 0. Es justamente a este propósito que presta atención a ciertos conjuntos de puntos en el intervalo **I**, a fin de determinar con precisión el conjunto de esos valores excepcionales. Cantor (1871) demuestra T2 —y por lo tanto la unicidad de la representación (1)— para el caso de que haya un número arbitrariamente grande pero finito de valores excepcionales de $x$ en cualquier intervalo finito dado. Cantor (1872) extiende las demostraciones anteriores a una familia de casos en que el intervalo finito **I** incluye un subconjunto *infinito* de puntos $x$ donde la secuencia (4) no converge o converge a un límite diferente de 0. En dicho trabajo —y al servicio de ese resultado— Cantor introduce su teoría de los números reales —que llama 'magnitudes numéricas' (*Zahlengrößen*)— y hace sus primeras consideraciones sobre conjuntos de puntos. Paso a explicarlas.

Cantor dice que "los números racionales constituyen el fundamento para la determinación del concepto más amplio de magnitud numérica". Los racionales forman "el dominio $A$" (Cantor, GA, p. 92). Sea $a_0$, $a_1$, . . . una secuencia infinita de racionales, tal que para cada número racional $\varepsilon > 0$ hay un entero positivo $N$, de suerte que $|a_m - a_n| < \varepsilon$ siempre que $N < m < n$. Diremos que $a_1$, $a_2$, . . . es una secuencia fundamental en el dominio $A$, y la designaremos $(a_i)$. (En vez de 'secuencia fundamental' suele decirse 'secuencia de Cauchy'.) Cantor asocia a cada secuencia fundamental en $A$ un índice sujeto a la condición siguiente: Si $(a_i)$ y $(b_i)$ son dos secuencias fundamentales en $A$ y para cada número racional $\varepsilon > 0$ hay un entero positivo

$N$ tal que $|a_n - b_n| < \varepsilon$ si $N < n$, se asocia un mismo índice a las secuencias $(a_i)$ y $(b_i)$. Diremos en tal caso que $(a_i)$ y $(b_i)$ son secuencias equivalentes.[5] Si $(a_i)$ y $(b_i)$ son dos secuencias fundamentales en $A$ asociadas a dos índices diferentes $\alpha$ y $\beta$, decimos que $\alpha < \beta$ si hay un número racional positivo $q$ y un entero positivo $N$ tales que $q < (b_n - a_n)$ si $N < n$. Cantor llama $B$ al dominio formado por los índices de las secuencias fundamentales de racionales. Claramente, la relación $<$ establece un orden lineal en $B$. Cantor define asimismo la adición y la multiplicación en $B$:

ADICIÓN: Sean $\alpha$ y $\beta$ los índices asociados respectivamente a las secuencias $(a_i)$ y $(b_i)$. Entonces, la suma $\alpha + \beta$ es el índice correspondiente a la secuencia $(a_i + b_i)$, formada sumando esas dos secuencias término a término.

MULTIPLICACIÓN: Sean $\alpha$ y $\beta$ los índices asociados respectivamente a las secuencias $(a_i)$ y $(b_i)$. Entonces, el producto $\alpha\beta$ es el índice correspondiente a la secuencia $(a_ib_i)$, formada multiplicando esas dos secuencias término a término.

En virtud de estas definiciones, los objetos que forman el dominio $B$ merecen llamarse "magnitudes numéricas". Las operaciones así definidas confieren a $B$ la estructura de un cuerpo† ordenado (aunque Cantor no se expresa así). Su próximo paso consiste en formar secuencias fundamentales con elementos tomados de $A$ y $B$. Ahora bien, el concepto de secuencia fundamental explicado arriba puede sin duda extenderse al dominio $B$ utilizando la adición y el orden definidos en éste;[6] pero sería preferible evitar la promiscuidad entre $A$ y $B$. Esto se logra reemplazando cada elemento $a$ de $A$ incluido en

[5] Como fácilmente se puede comprobar, la relación entre $(a_i)$ y $(b_i)$ es una genuina relación de equivalencia, puesto que es una relación simétrica, transitiva y reflexiva. Cabe, pues, entender que el índice asociado a cada secuencia fundamental no es otra cosa que la clase de equivalencia a la que pertenece en virtud de esta relación.

[6] La adición determina el 0 en $B$, a saber, aquel objeto cuya suma con cualquier otro es igual a este último. Los demás objetos de $B$ quedan entonces clasificados en positivos y negativos, según sean mayores o menores que el 0. Si $\alpha$ y $\beta$ están en $B$, $\alpha - \beta$ designa la suma de $\alpha$ y el único objeto de $B$ cuya suma con $\beta$ es igual a 0. Una secuencia fundamental en $B$ puede entonces definirse en los mismos términos utilizados arriba para definir una secuencia fundamental en $A$, con una sola diferencia: $\varepsilon > 0$ debe ser un objeto de $B$, no de $A$.

una secuencia promiscua por el elemento de $B$ correspondiente a la secuencia constante $(a, a, \ldots)$ — esto es, a la secuencia $(a_k)$ tal que $a_k = a$ para cada entero positivo $k$. Mediante un procedimiento análogo al seguido para formar el dominio $B$, Cantor forma un dominio $C$ de índices asociados a las secuencias fundamentales en $B$, un dominio $D$ de índices asociados a las secuencias fundamentales en $C$, etc. La equivalencia de secuencias fundamentales, así como el orden, la adición y la multiplicación de los índices pertinentes se definen en cada nuevo dominio de una manera similar a la indicada en la construcción de $B$.

Es claro que el dominio $B$, aunque dotado de una estructura algebraica parecida a la de $A$, es esencialmente más rico que éste. En efecto, si bien cada elemento de $A$ corresponde, del modo explicado, a un elemento único de $B$, hay elementos de $B$ que no corresponden de este modo a ningún elemento de $A$. Por ejemplo, si $(a_i)$ es una secuencia en $A$ tal que la secuencia de productos $(a_i a_i)$ es equivalente a la secuencia constante $(2, 2, \ldots)$, no hay un número racional $q$ tal que $(a_i)$ sea equivalente a la secuencia constante $(q, q, \ldots)$ y por lo tanto el índice de la secuencia $(a_i)$ no corresponde del modo antedicho a ningún elemento de $A$. En cambio, toda secuencia fundamental $(\alpha_i)$ de elementos de $B$ equivale a alguna secuencia constante $(\alpha, \alpha, \ldots)$ en $B$, y otro tanto puede decirse de las secuencias fundamentales en $C$, $D$, etc. Por esta razón, los dominios $C$, $D$, ... son todos estructuralmente idénticos a $B$. Sea $K$ uno cualquiera de estos dominios. Entonces hay una aplicación biyectiva única de $B$ en $K$ que respeta la relación $<$ y todas las relaciones determinadas por la adición y la multiplicación.[7] Dicha aplicación es lo que se llama un *isomorfismo*[†] (porque retrata fielmente una estructura en la otra) *canónico* (porque es único en su género). Salvo una particular afición a la recurrencia infinita, no veo qué pueda haber inducido a Cantor a reconocer los dominios $C$, $D$, ... como diferentes de $B$.

Por último, Cantor coordina los elementos del dominio $B$ con los puntos de una recta (parametrizada) cualquiera. Recuerda que, una vez que se ha fijado el punto 0, las direcciones positiva y negativa desde el 0 y la unidad de distancia, cada punto $p$ queda perfectamente individualizado por su dis-

[7] Sea $\varphi: B \to K$ la aplicación en cuestión. Entonces, para cada $\alpha, \beta \in B$, $\varphi(\alpha) < \varphi(\beta)$ si y sólo si $\alpha < \beta$; $\varphi(\alpha + \beta) = \varphi(\alpha) + \varphi(\beta)$, y $\varphi(\alpha\beta) = \varphi(\alpha)\varphi(\beta)$. En particular, $\varphi(0)$ es el 0 de $K$ y $\varphi(1)$ es el 1 de $K$.

tancia (orientada) al punto 0. Si esa distancia es un número racional *a*, el punto *p* queda asociado unívocamente a ese elemento de *A*. (Y también, por ende, al elemento único de *B* que es el índice de la secuencia constante (*a*, *a*, . . . )). Pero si la distancia orientada de *p* al 0 no es un número racional, siempre habrá una secuencia fundamental de racionales ($a_i$) tal que (i) para cada entero positivo *k* hay un punto $p_k$ a distancia $a_k$ del 0, y (ii) todo intervalo centrado en *p*, por pequeño que sea, encierra infinitos puntos de la secuencia ($p_i$). Cantor resume esto diciendo que la distancia de *p* al punto 0 es igual a *b*, donde *b* es el índice de la secuencia ($a_i$). De este modo, a cada punto de la recta corresponde una y sólo una "magnitud numérica" en *B*. Cantor reconoce que no es posible demostrar que exista la correspondencia inversa, en virtud de la cual "a cada magnitud numérica le pertenece un determinado punto de la recta, cuya coordenada es igual a esa magnitud numérica, en el sentido aquí explicado". Por eso la acepta simplemente como "axioma". En virtud de este axioma, "las magnitudes numéricas obtienen una cierta objetividad, de la cual, sin embargo, son enteramente independientes" (Cantor, GA, p. 97).

Establecida así —por decreto— la correspondencia entre las "magnitudes numéricas" (números reales) y los puntos de la recta, Cantor advierte que en lo sucesivo cuando hable de puntos se referirá a las magnitudes numéricas correspondientes. En aras de la brevedad llamará "conjunto de valores" (*Wertmenge*) a una multitud dada de tales magnitudes, y "conjunto de puntos" (*Punktmenge*) a la correspondiente multitud de puntos. Si *P* es un conjunto cualquiera de puntos contenidos en un intervalo finito, *P* determina otros conjuntos de puntos, que Cantor llama "conjuntos derivados" (*abgeleitete Punktmengen*). Se definen así. Un entorno de un punto *p* es un intervalo que contiene a *p*. Si cada entorno de *p* contiene infinitos puntos del conjunto *P*, diremos que *p* es un *punto límite* (*Grenzpunkt* — también se dice 'punto de acumulación') de *P*. El primer conjunto derivado $P'$ es el conjunto de todos los puntos límite del conjunto *P*. El ($\nu+1$)-ésimo conjunto derivado $P^{(\nu+1)}$ es el conjunto de todos los puntos límite del $\nu$-ésimo conjunto derivado $P^{(\nu)}$. Por

8 Sea dicho de paso, aquí tenemos un ejemplo de conjunto unitario, propuesto por Cantor mismo. El texto, traducido, dice así: "Si el conjunto *P* consiste de los puntos correspondientes a las abscisas 1, 1/2, 1/3, . . . ,1/*n*, . . . el conjunto $P'$ consiste del solo punto 0 y no tiene un conjunto derivado." (Cantor, GA, p. 98).

ejemplo, si $P$ es el conjunto de puntos a distancia 1, 1/2, 1/3, 1/4, . . . del 0, el primer conjunto derivado $P'$ es el conjunto $\{0\}$.[8] Evidentemente, si el $\nu$-ésimo conjunto derivado $P^{(\nu)}$ de un conjunto de puntos $P$ contiene sólo un número finito de puntos, su $(\nu+1)$-ésimo conjunto derivado $P^{(\nu+1)}$ no contiene ninguno. Hoy diríamos que en tal caso $P^{(\nu+1)} = \varnothing$, pero lo que dice Cantor es que si $P^{(\nu)}$ contiene sólo un número finito de puntos, entonces $P^{(\nu)}$ *no tiene* conjunto derivado. El conjunto original $P$ es entonces lo que Cantor llama un "conjunto de puntos de la $\nu$-ésima especie" (*Punktmenge* $\nu^{ter}$ *Art*).

Recurriendo a estos conceptos, Cantor demuestra que la proposición T2 sigue siendo válida aunque sus condiciones no se cumplan en un conjunto de puntos de la $\nu$-ésima especie incluido en el intervalo **I**, cualquiera que sea el entero positivo $\nu$ (siempre, claro está, que esas condiciones estén satisfechas en el resto de **I**). Ello implica el siguiente teorema sobre la unicidad de la representación (1):

> Si la serie trigonométrica
>
> $$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty}(a_n \cos nx + b_n \operatorname{sen} nx)$$
>
> representa a la función real $f$ en todos los puntos $x$ del conjunto $\mathbf{I} \setminus P$, donde **I** es un intervalo finito de la recta real y $P \subset \mathbf{I}$ es un conjunto de puntos tal que su conjunto derivado $P^{(n)} = \varnothing$ para algún entero positivo $n$, entonces dicha representación es única.

## 1.4 DIVERSOS INFINITOS

Como veremos en el Capítulo 1.5, con sus reflexiones sobre los conjuntos derivados Cantor dio un primer paso decisivo para la construcción de la aritmética del infinito expuesta en los trabajos “Sobre variedades lineales infinitas de puntos” de 1879–1884. Pero antes explicaré ciertos hallazgos sobre conjuntos infinitos que Cantor publicó en 1874 y 1878 y que contribuyeron poderosamente a motivar sus estudios posteriores. Se refieren a lo que Cantor llama la “potencia” (*Mächtigkeit*), esto es, la numerosidad de un conjunto, que es la única propiedad que lo caracteriza si se hace caso omiso de su estructura y de la naturaleza de sus miembros.[1] Decimos que dos conjuntos $a$ y $b$ son *equinumerosos*, o que tienen la misma potencia, si hay una aplicación† biyectiva de $a$ sobre $b$. En cambio, si hay una aplicación inyectiva de $a$ en $b$, pero no hay una aplicación biyectiva de $a$ sobre $b$, diremos que $b$ es más numeroso que $a$, o que su potencia es mayor que la de $a$. Evidentemente, estas definiciones prescinden de la índole de $a$ y $b$, y de sus respectivos elementos. Utilizando estos conceptos, Cantor (1874) demuestra las proposiciones siguientes:

[1] Tras explicar con admirable claridad y concisión la médula de las investigaciones de Cantor sobre la representación única de funciones reales mediante series trigonométricas, Maddy indica sugestivamente de qué modo el descubrimiento del teorema enunciado al final del Capitulo 1.3 pudo conducirle a reflexionar sobre la numerosidad de los conjuntos infinitos. Una vez que definió los reales en términos de secuencias fundamentales, Cantor pudo extender su teorema de unicidad de 1870 a funciones que convergen sólo en el complemento de un conjunto infinito, aunque excepcional, de puntos.

> Pero ¡qué curioso conjunto de puntos era éste! Finito y bastante complejo, y sin embargo en cierto modo tan pequeño o tan bien portado con respecto a la totalidad de los reales, que resultaba inofensivo. Al parecer, esto puso a Cantor a pensar de cómo era que los conjuntos continuos, tales como los reales, se relacionan con conjuntos infinitos discretos y al parecer más pequeños, tales como los números naturales.
>
> (Maddy 1990, p. 108)

(I) El conjunto de todos los números algebraicos y el conjunto de los enteros positivos {1, 2,...} son equinumerosos.

(II) El conjunto de los números reales comprendidos en un intervalo finito cualquiera es más numeroso que el conjunto de los enteros positivos.

Cantor (1878) demuestra que:

(III) El conjunto de los puntos contenidos en un segmento recto es equinumeroso con el conjunto de los puntos contenidos en un cuadrado o en un cubo o, en general, en cualquier hipercubo de dimensión arbitraria $n$.

Se llama número algebraico a cualquier número real que sea una solución de una ecuación con coeficientes enteros, de la forma:

$$\alpha_0 + \alpha_1 x + \alpha_2 x^2 + \ldots + \alpha_n x^n = 0 \qquad (1)$$

El índice más alto para el cual $\alpha_n \neq 0$ se llama el grado de la ecuación (1). Una ecuación de grado $n$ admite a lo sumo $n$ soluciones reales. Para que cada solución corresponda a una sola ecuación de la forma (1), Cantor requiere además que $\alpha_0$ sea positivo y que todos los coeficientes sean primos entre sí (es decir, que no tengan un divisor común). Digamos que una ecuación que reúne estas características es una ecuación simplificada. Cantor asigna a cada número algebraico un entero positivo que llama su altura (*Höhe*) y que determina así: Sea $u$ una solución de una ecuación simplificada de grado $n$ con coeficientes $\alpha_0, \alpha_1, \ldots, \alpha_n$. Entonces la altura de $u$ es igual a $n - 1 + |\alpha_0| + |\alpha_1| + \ldots + |\alpha_n|$. Se comprueba fácilmente que no hay más que una cantidad finita de números algebraicos de una misma altura. Cantor observa que hay sólo uno con altura 1, dos con altura 2, cuatro con altura 3. Digamos que el número algebraico $u$ precede al número algebraico $v$ si $u$ tiene menor altura que $v$, o, en el caso de que tengan la misma altura, si $u$ es menor que $v$. Evidentemente, los números algebraicos ordenados de esta manera forman una secuencia infinita, $u_1, u_2, \ldots$, en correspondencia biunívoca con los enteros positivos. Este resultado llama la atención, ya que en cada entorno

de cualquier número real $p$ hay infinitos números algebraicos.

Para probar que todo intervalo real $(\alpha, \beta)$ es más numeroso que el conjunto de los enteros positivos, Cantor considera una secuencia infinita arbitraria $\sigma$ de números reales, todos diferentes, $p_1, p_2, \ldots$, y muestra que en cualquier intervalo $(\alpha, \beta)$ hay por lo menos un número real que no pertenece a esa secuencia. Esto significa que ninguna aplicación inyectiva del conjunto de los enteros positivos en el conjunto de los reales cubre todo el intervalo $(\alpha, \beta)$. Vamos a recorrer en orden la secuencia $\sigma$ hasta encontrar dos números contenidos en $(\alpha, \beta)$. Desígnemoslos $\alpha_1$ y $\beta_1$, de tal modo que $\alpha_1 < \beta_1$. (Notamos que si no hubiera más que un número $p_r$ en $\sigma$ contenido en $(\alpha, \beta)$, cualquier elemento del interior de $(\alpha, p_r)$ sería un número real que no pertenece a $\sigma$). Sigamos recorriendo la secuencia $\sigma$, hasta hallar dos números contenidos en el intervalo $(\alpha_1, \beta_1)$. Llamémoslos $\alpha_2$ y $\beta_2$, de modo que $\alpha_2 < \beta_2$. Esta operación se repetirá tantas veces como dé un resultado, designándose, en general, con $\alpha_{n+1}$ y $\beta_{n+1}$ a los dos primeros miembros de la secuencia $\sigma$ contenidos en el intervalo $(\alpha_n, \beta_n)$ y tales que $\alpha_{n+1} < \beta_{n+1}$. Hay dos posibilidades: o bien (**P1**) hay un entero positivo $k$ tal que el interior de $(\alpha_k, \beta_k)$ no contiene dos números de $\sigma$, en cuyo caso ninguno de los infinitos números reales mayores que $\alpha_k$ y menores que $\beta_k$ pertenece a $\sigma$; o bien (**P2**), no importa cuán grande sea $k$, el intervalo $(\alpha_k, \beta_k)$ siempre contiene por lo menos dos números de $\sigma$, en cuyo caso la secuencia infinita creciente $\alpha_1, \alpha_2, \ldots$, acotada arriba por los $\beta_i$, converge a un límite $\alpha^*$, y la secuencia infinita decreciente $\beta_1, \beta_2, \ldots$, acotada abajo por los $\alpha_i$, converge a un límite $\beta^*$. Si $\alpha^* < \beta^*$, cualquier número contenido en el intervalo $(\alpha^*, \beta^*)$ es un número real que no pertenece a $\sigma$. Si $\alpha^* = \beta^*$, entonces $\alpha^*$ no pertenece a $\sigma$, puesto que, si $\alpha^*$ perteneciera a $\sigma$, tendríamos que $\alpha^* = p_r$ para cierto entero positivo $r$ y las secuencias $\alpha_1, \alpha_2, \ldots$ y $\beta_1, \beta_2, \ldots$ son subsecuencias de $\sigma$ definidas de tal modo que $p_r$ no puede estar contenido en el intervalo $(\alpha_r, \beta_r)$ (en el mejor de los casos, $\{\alpha_1, \beta_1\} = \{p_1, p_2\}$ y el primer miembro de $\sigma$ que puede estar contenido en $(\alpha_1, \beta_1)$ es $p_3$); en cambio, $\alpha^*$ está contenido en $(\alpha_r, \beta_r)$, cualquiera que sea el entero positivo $r$. Por lo tanto, existe al menos un número real en el intervalo $(\alpha, \beta)$ que no pertenece a la secuencia $\sigma$.

Cantor (1874) habla de los enteros positivos, de los números algebraicos y de los números reales en cierto intervalo finito como si cada una de estas

clases infinitas de números estuviera presente en su totalidad.[2] El supuesto de que la investigación matemática se ocupa con conjuntos bien definidos que constan de infinitos elementos está, por cierto, implícito ya en su teoría de los números reales, según la cual cada uno de estos corresponde a una secuencia fundamental infinita bien definida de números racionales; o, para ser más exacto, a una clase infinita de tales secuencias.[3] Este supuesto es contrario a la concepción tradicional, heredada de Aristóteles, según la cual la infinitud de los números y otras clases de objetos matemáticos —que se manifiesta en la posibilidad de producir otros nuevos, interminablemente, cada vez que se ha fijado una cierta cantidad de ellos— excluye el que pueda concebírselos como formando una totalidad acabada. Más adelante veremos como Cantor defiende en sus escritos filosóficos la actualidad del infinito matemático, contra esta idea aristotélica de un infinito meramente potencial. Pero conviene examinar desde ya en qué medida la concepción cantoriana del infinito afecta no sólo el *enunciado* sino también la *demostración* de las proposiciones I y II.

Pienso que la demostración de la proposición I no presupone un infinito actual. Cantor propone un método para enumerar los números algebraicos de modo que ninguno se omita. Para ello no es menester que los enteros positivos formen una totalidad dada de antemano. Basta irlos tomando en orden, uno a uno, de una fuente virtualmente inagotable, y asignarlos a los números algebraicos conforme a una regla que asegure que ninguno de éstos será pasado por alto. Esto es lo que se logra con el método propuesto. Supongamos que cuando enumeramos cierto número $a$ cierto número algebraico $b$ no ha sido enumerado todavía. En tal caso, o bien $b$ tiene mayor altura que $a$, o, si tiene la misma altura, es simplemente mayor que $a$. En ambos casos, le corresponde ser enumerado después que $a$ y, por consiguiente,

[2] Cantor, GA, pp. 115, 116, menciona expresamente “der Inbegriff aller ganzen positiven Zahlen”, “die reellen algebraischen Zahlen in ihrer Gesamtheit” y “die sämtlichen reellen Zahlen, welche $\geq 0$ und $\leq 1$ sind”.

[3] Adviértase que si $\sigma$ es una secuencia fundamental y se modifican de cualquier manera los $n$ primeros términos de $\sigma$, donde $n$ es un entero determinado, tan grande como se desee, se obtiene una secuencia fundamental $\sigma'$ equivalente a $\sigma$. Así, la identidad de cada número real, según la teoría de Cantor, depende de la última parte —infinita— de la secuencia fundamental que lo representa (esto es, de lo que resta de la secuencia después de remover los $n$ primeros términos, para $n$ fijo, pero arbitrario).

su ausencia de entre los números algebraicos enumerados antes que $a$ no constituye una omisión.

En cambio, parece que la demostración de la proposición II depende esencialmente de que el conjunto de los enteros positivos esté ya dado. Se trata de mostrar que, como quiera que este conjunto se inyecte en el dominio de los reales, la imagen de la inyección excluye por lo menos un número real de un determinado intervalo.[4] Para probar la proposición II Cantor considera, como vimos, una inyección cualquiera $\sigma$, cuya imagen realiza una de las alternativas que arriba llamé **P1** y **P2**. Ahora bien, sólo se puede decidir que estamos frente a la alternativa **P1** si hemos recorrido todos los valores de $\sigma$. De otro modo ¿cómo se podría establecer que a lo sumo hay sólo uno de esos valores dentro de cierto intervalo $(\alpha_k, \beta_k)$, determinado como se explicó? Sin embargo, sería errado creer que la prueba de la proposición II requiere que se sepa cuál de las dos alternativas se cumple. Sólo hace falta saber que, cualquiera que sea la aplicación inyectiva $\sigma$ de los enteros positivos en los reales, tiene que cumplirse una de las dos. La argumentación de Cantor va dirigida a mostrar que, en ambos casos, hay dentro del intervalo $(\alpha, \beta)$ por lo menos un número real que no es un valor de $\sigma$. Como **P1** es la negación de **P2,** el principio lógico del "tercero excluido" implica que una tercera alternativa es imposible.[5] (Según L. E. J. Brouwer, la matemática cantoriana es inaceptable justamente porque dicho principio lógico no vale en los dominios infinitos; cf. Apéndice XIX).

La demostración de la proposición III publicada por Cantor en 1878 es demasiado larga para reproducirla aquí.[6] Pero una carta suya a Dedekind de 1877 contiene otra demostración más sencilla que voy a parafrasear. Inyectaremos el interior de un cuadrado $Q$ en el interior de un segmento $S$. Con este propósito le asignamos a cada punto en éste último un número real $x$

4 Si el conjunto de los enteros positivos existe, no cabe duda de que hay por lo menos una aplicación inyectiva del mismo en el dominio de los reales, a saber, la inclusión canónica que asigna a cada entero positivo $z$ el número real correspondiente a la secuencia constante $(z, z, \ldots)$.

5 **P1** se cumple si para todo entero positivo $k$ hay dos valores de la secuencia $\sigma$ en el intervalo $(\alpha_k, \beta_k)$. **P2** se cumple si hay un entero positivo $k$ tal que en el intervalo $(\alpha_k, \beta_k)$ no hay dos valores de $\sigma$.

6 El texto alemán original aparece en Cantor, GA, pp. 119–133. Dauben 1979, pp. 58–66, da una buena paráfrasis en inglés.

en el intervalo (0,1) y a cada punto en *Q* un par de coordenadas $\langle y,z \rangle$ del mismo intervalo. Cualquier número de ese intervalo puede expresarse de una sola manera como un decimal infinito $0{,}\alpha_1\alpha_2\alpha_3 \ldots$, donde cada $\alpha_i$ es un dígito y para cada entero positivo *k* hay otro entero *h* tal que $\alpha_{k+h} > 0$. Algunos de esos números pueden además representarse mediante decimales finitos —por ejemplo, 0,273999 . . . = 0,274— pero excluiremos esta forma de expresión para evitar duplicaciones. Nuestra aplicación inyectiva asigna al punto de *Q* con coordenadas $y = 0{,}\alpha_1\alpha_2\alpha_3 \ldots$ y $z = 0{,}\beta_1\beta_2\beta_3 \ldots$ el punto de *S* correspondiente al número $x = 0{,}\alpha_1\beta_1\alpha_2\beta_2\alpha_3\beta_3 \ldots$ Dedekind objetó a Cantor que esta aplicación no es biyectiva. En efecto, su imagen no incluye ningún punto de *S* cuya coordenada se exprese mediante un decimal infinito de la forma $0{,}\xi_1\xi_2 \ldots \xi_k 0 \xi_{k+2} 0 \xi_{k+4} 0 \ldots$ (en el cual, indefinidamente, a partir de cierto dígito, hay, dígito por medio, un cero). Cantor aceptó esta crítica y elaboró otra demostración más engorrosa. Sin embargo, a la luz de un teorema demostrado (independientemente) por E. Schröder y F. Bernstein casi veinte años más tarde, el argumento original de Cantor resulta suficiente para probar la proposición III. Según ese teorema, si un conjunto *a* es inyectable en un conjunto *b* y *b* es inyectable en *a*, *a* y *b* son equinumerosos.[7] Como es claro que el segmento *S* se puede inyectar en el cuadrado *Q* (por ejemplo, mediante la aplicación $x \mapsto \langle x,x \rangle$), la existencia de una aplicación inyectiva de *Q* en *S* prueba que tienen la misma potencia. El método seguido para definir la inyección de *Q* en *S* se deja extender fácilmente al caso en que *Q* sea un hipervolumen de cualquier dimensión.

La publicación de la proposición III vino a agravar la llamada crisis de la intuición matemática en el último tercio del siglo XIX.[8] Intuitivamente parece haber una diferencia irreductible entre un recta y una superficie o un volumen. Los matemáticos conceptualizaban esa diferencia relacionándola con el número de coordenadas requerido para identificar cada punto de estos “continuos” (una, dos y tres, respectivamente). Por esta vía, se hacía fácil concebir los “hiperespacios” de más de tres dimensiones: era cosa de incre-

7 Las primeras demostraciones del teorema de Schröder y Bernstein solían aducir premisas más fuertes de lo necesario. Se hallará una demostración más económica en Levy 1980, pp. 85s.

8 Se agrega a la “curva sin dirección” de Weierstraß, una función real continua pero no diferenciable. Precede a la “curva que llena una superficie” de Peano, aplicación continua de un segmento recto sobre un cuadrado.

mentar el número de coordenadas asignadas a sus puntos. Pero la proposición III muestra que basta con una sola coordenada para identificar los puntos de un continuo de cualquier número de dimensiones. Dedekind sugirió que sólo podía haber biyecciones *continuas* entre dominios equidimensionales. Pero trascurrirán más de 30 años antes de que Brouwer (1911, 1913) demuestre este teorema.[9] Para ello hubo que encontrar primero una caracterización viable del concepto de número de dimensiones. (Cf. D. M. Johnson 1978/81).

Las proposiciones I y II distinguen dos clases de infinito: el infinito *denumerable* de los enteros positivos y algebraicos, y el infinito mayor de los números reales. Cantor pronto conjurará una sucesión hiperinfinita de infinitos, cada cual más numeroso que los anteriores. Junto a su demostración de que todos los continuos son equinumerosos publicó una conjetura que luego intentará sin éxito demostrar: “Entre las variedades lineales infinitas no se encuentran más que dos potencias” (Cantor, GA, p. 133), a saber, la de los enteros y la de los reales. En otras palabras: un continuo $C$ nunca incluirá un subconjunto $D$ que sea a la vez menos numeroso que $C$ y más numeroso que los enteros positivos. Esta es la célebre *Hipótesis del Continuo* de Cantor.

9 Cantor (1878) subraya que su rotulación de cada punto del plano mediante una sola coordenada supone que la parametrización no sea continua, pero expresamente deja abierta la cuestión que será resuelta por Brouwer (GA, p. 121).

Al comienzo del Nº 5 de la serie "Sobre variedades lineales infinitas de puntos", publicado en 1883, Cantor declara que sus investigaciones acerca de ese tema han alcanzado un punto desde el cual no puede dar ni un paso más con naturalidad (*zwanglos*) si no extiende el concepto de número entero más allá de sus límites acostumbrados. Esta es su justificación o, "si fuese menester", su excusa para introducir en ese escrito "ideas aparentemente extrañas". Como explica enseguida, se trata de ampliar o, más bien, de continuar la serie de los números enteros "más allá del infinito" (GA, p. 165).

Cantor sabía muy bien que la matemática tradicional —lo que habrían llamado "matemática clásica" en 1880 si este modo de expresarse hubiera sido corriente a la sazón— admitía el infinito sólo como una potencialidad inalcanzable y rechazaba de plano el infinito actual.

> Me parece que el infinito matemático, en la medida en que ha encontrado en la ciencia hasta la fecha una aplicación justificada y provechosa, se presenta ante todo en la acepción de una cantidad variable que, o bien crece más allá de todo límite, o bien decrece hasta ser arbitrariamente pequeña, pero que siempre sigue siendo *finita*. Llamo a este infinito, el *infinito impropio* (*das Uneigentlich-unendliche*).
>
> (Cantor, GA, p. 165)

Cantor patrocina, en cambio, una concepción del infinito matemático como algo completamente determinado. Aduce como ejemplo el "punto en el infinito" utilizado ya entonces en la teoría de la funciones de una variable compleja. Pero, nos advierte,

> mientras el punto en el infinito del plano complejo se yergue solo frente a todos los puntos del dominio finito, aquí obtendremos no tan sólo un número entero infinito único, sino una secuencia infinita de tales números, que se distinguen bien unos de otros y sostienen relaciones aritméticas regulares entre sí y con los números enteros finitos.
>
> (Cantor, GA, p. 165)

El avance al transfinito procede en dos frentes. Ya conocemos uno. Vimos como Cantor demostraba que el conjunto $\mathbb{R}$ de los números reales (o de los puntos de la recta) incluye subconjuntos tales como el conjunto de los números enteros y el de los números algebraicos (respectivamente, los puntos con coordenadas enteras y con coordenadas algebraicas), que son infinitos pero menos numerosos que $\mathbb{R}$. Según esto, en $\mathbb{R}$ hay por lo menos dos modos o niveles diferentes de infinitud. Cantor (1890/91) mostrará que, dado un conjunto cualquiera $K$, el conjunto de sus partes $\mathcal{P}K$ es siempre más numeroso que $K$. Resulta entonces que, si sobreentendemos con Cantor que, dado un conjunto, también está dado el conjunto de sus partes, basta que haya *una* numerosidad infinita, para que haya *infinitas otras* mayores que ella. Cantor elaborará una aritmética de tales numerosidades, con reglas precisas para la adición, la multiplicación, la exponenciación. Esta línea de avance tiene, con todo, un inconveniente: las numerosidades generadas por la operación $K \mapsto \mathcal{P}K$ se suceden, sí, de menor a mayor, pero no hay cómo saber si no existen numerosidades intermedias entre la numerosidad de un conjunto $K$ y la numerosidad de $\mathcal{P}K$. Pero Cantor avanzará al transfinito también por otro frente en el cual ese inconveniente no se presenta (al menos, si se acepta una hipótesis que él juzga verdadera).

Cuando comparamos la numerosidad de dos conjuntos infinitos $M_1$ y $M_2$ con los métodos de la Capitulo 1.4 lo que hacemos puede describirse así: Los elementos del conjunto $M_1$ se emplean como índices para enumerar los elementos del conjunto $M_2$; si alcanzan para todos, $M_1$ no es menos numeroso que $M_2$; pero si, como quiera que se marquen los elementos de $M_2$ con índices tomados de $M_1$, necesariamente queda alguno sin marcar, es claro que $M_2$ es más numeroso que $M_1$. Si luego queremos comparar la numerosidad de $M_2$ con la de un tercer conjunto $M_3$ no podemos usar a $M_1$ como conjunto de índices, sino que tenemos que valernos de los elementos de $M_2$ o $M_3$. No es así como comparamos ordinariamente la numerosidad de conjuntos finitos. Antes bien, la serie de los enteros positivos, ordenada de menor a mayor, nos sirve aquí como único sistema de índices, con el cual enumeramos los elementos de cada uno de los conjuntos que queremos comparar. Por esta vía, junto con averiguar que uno de los conjuntos así enumerados es más numeroso que otro, se establece exactamente cuántas numerosidades intermedias hay entre las de esos dos conjuntos. La segunda línea de avance seguida por Cantor provee una continuación transfinita de la sucesión ordenada de los enteros.

La idea de una tal continuación surge con toda naturalidad del estudio de los conjuntos derivados de un conjunto de puntos. Cuando hablé de ellos en el Capítulo 1.3 me referí expresamente sólo a conjuntos de la $n$-ésima especie, cuyo $n$-ésimo conjunto derivado es finito, de suerte que el conjunto derivado $(n + 1)$-ésimo está vacío. Pero la recta incluye por cierto más de un conjunto de puntos $P$ tal que, para cualquier entero positivo $n$, el conjunto derivado $P^{(n)} \neq \varnothing$. Tenemos además que, si bien $P^{(1)}$ generalmente no es una parte de $P$ (por ejemplo, si $P$ es el conjunto de los puntos con coordenadas algebraicas, $P^{(1)} = \mathbb{R}$), en cambio, $P^{(n+1)} \subseteq P^{(n)}$ para todo $n \geq 1$. Por ende, $P^{(n)} \neq \varnothing$, la intersección infinita $\bigcap_{i=1}^{\infty} P^{(i)} \neq \varnothing$ . Cantor llama a esta intersección "la derivación de orden $\infty$" y la designa con el símbolo $P^{(\infty)}$. Considera en seguida la serie de sus conjuntos derivados $P^{(\infty+1)}$, $P^{(\infty+2)}$, . . . También $P^{(\infty)}$ tiene su derivación de orden $\infty$, generalmente no vacía. Cantor la designa con $P^{(2\infty)}$. "Prosiguiendo con estas construcciones conceptuales uno llega a derivaciones que es consecuente designar con $P^{(n\infty+m)}$, donde $n$ y $m$ son números enteros positivos" (GA, p.147). La intersección $\bigcap_{k=1}^{\infty} P^{(k\infty)}$ debe llamarse entonces $P^{(\infty^2)}$. El próximo paso natural es formar derivaciones cuyo orden está dado por un polinomio en $\infty$ de grado $n$ $(n > 1)$. La intersección infinita de tales derivaciones es, lógicamente, $P^{(\infty^\infty)}$.

> Avanzando de modo consecuente se ganan sucesivamente los ulteriores conceptos:
>
> $$P^{(\infty^\infty)},\ P^{(\infty^{\infty+1})},\ P^{(\infty^{\infty+n})},\ P^{(\infty^{n^\infty})},\ P^{(\infty^{\infty^n})},\ P^{(\infty^{\infty^\infty})},\ \text{etc.};$$
>
> vemos aquí una generación dialéctica de conceptos que conduce cada vez más lejos, manteniéndose libre de toda arbitrariedad, como algo en sí mismo necesario y consecuente.
>
> (Cantor, GA, p. 148)

Como el símbolo $\infty$ se utiliza tradicionalmente en matemáticas para significar el infinito potencial a que tiende una cantidad variable no acotada, Cantor lo reemplazó en 1883, en su nuevo uso, por el símbolo $\omega$, que emplearemos en adelante (GA, p. 195n.). El objeto $\omega$ surge, pues, como el primer índice transfinito en la serie de las derivaciones de un conjunto de puntos. Pero, al igual que los índices finitos, $\omega$ y sus sucesores pueden separarse de esta

función y usarse para enumerar los elementos de otros conjuntos. Tomados así en abstracto constituyen, según Cantor, una continuación natural de la serie de los enteros positivos, suficiente para medir “todas las diversas numerosidades, en sucesión ascendente, presentes en la naturaleza corpórea y espiritual” (GA, p. 199).

Antes de hacer esta asombrosa aseveración, Cantor explica ciertos conceptos que precisan su sentido y alcance. Para facilitar la exposición adoptaré ciertas convenciones. Llamaré ‘ordinales’ a los elementos —finitos y transfinitos— de la serie cantoriana de enteros (este término, de uso corriente, corresponde al alemán *Ordnungszahlen* que Cantor emplea en sus publicaciones desde 1887; GA, p. 388). Siguiendo la práctica actual, entenderé que el primer ordinal es el 0, no el 1 (cf. Cantor, GA, p. 445). Para denotar ordinales en general usaré las minúsculas del alfabeto griego, aunque recurriré a la cursiva minúscula de nuestro alfabeto cuando quiera referirme específicamente a un ordinal finito, esto es, a un predecesor de $\omega$. La expresión ‘$\alpha < \beta$’ significa que $\alpha$ precede a $\beta$ en la sucesión de los ordinales. Puede leérsela ‘$\alpha$ es menor que $\beta$’ o ‘$\beta$ es mayor que $\alpha$’. Por ahora —hasta el Capítulo 1.8— designaré con $[\alpha]$ al conjunto de todos los ordinales menores que un cierto ordinal $\alpha$. (Hoy día no se necesita un símbolo especial para designarlo, desde que von Neumann caracterizó los ordinales de tal modo que $\alpha = [\alpha] = \{\xi : \xi < \alpha\}$, *por definición*; véase la Sección 1.8.4.) Tenemos, entonces, que $[\alpha] \subseteq [\beta]$ si $\alpha < \beta$. Si los elementos de un conjunto $K$ se dejan enumerar exhaustivamente utilizando en orden todos los predecesores del ordinal $\kappa$, diré que $\kappa$ es un *enumerador* de $K$. (Con este término artificial traduzco aquí el vocablo alemán *Anzahl*, que Cantor usa en esta acepción).

Obviamente, si $\kappa$ es un *enumerador* de $K$, los conjuntos $K$ y $[\kappa]$ son equinumerosos. Si $K$ es un conjunto finito, es decir, si algún $k < \omega$ es un enumerador de $K$, entonces, no importa cómo $K$ se enumere, su enumerador es siempre el mismo. Cabe afirmar, pues, que $k$ mide la numerosidad de $K$ o, como decimos ordinariamente en castellano, que $k$ es el *número* de los elementos de $K$. (Así, por ejemplo, 4 es el número de los evangelios porque, en cualquier orden que se los tome, se los puede enumerar con los ordinales del conjunto $[4] = \{0, 1, 2, 3\}$). Pero si $K$ es infinito, admitirá distintos enumeradores según el orden en que se enumeren sus elementos. Por ejemplo, si $K = [\omega]$, su enumerador es $\omega$ si los elementos de $K$ se toman en el orden habitual; pero es $\omega^2$ si se los enumera en el orden siguiente: primero

el 0 y el 1, seguidos de todos los números primos, 2, 3, 5, 7, . . . ; luego los cuadrados de los números primos; luego sus respectivos cubos; . . . ; luego la *n*-ésima potencia de cada primo, en orden ascendente; . . . luego . . . Por eso, Cantor cree necesario distinguir entre el número (*Zahl*) y el enumerador (*Anzahl*) de un conjunto.[1]

Evidentemente, una vez determinado el orden en que debe enumerarse un conjunto $K$, se fija de modo inequívoco su enumerador. De hecho, la enumeración exhaustiva de $K$ por el conjunto ordenado $[\kappa]$ introduce (o presupone) en $K$ un orden particularísimo, a saber, el orden de los predecesores de $\kappa$ en la sucesión de los ordinales. Dicho orden tiene ciertas características generales, cualquiera que sea $\kappa$. Desde luego, reúne todos los requisitos de lo que se llama un *orden lineal*, a saber, (i) si $\alpha, \beta \in [\kappa]$, se cumple en todo caso una y sólo una de estas tres alternativas: $\alpha < \beta$, o $\beta < \alpha$, o $\alpha = \beta$; (ii) si $\alpha, \beta, \gamma \in \kappa$, $\alpha < \beta$ y $\beta < \gamma$, entonces $\alpha < \gamma$.[2] Pero no todo orden lineal concuerda con el orden de sucesión de los elementos de $[\kappa]$, para algún ordinal $\kappa$. Considérese, por ejemplo, el conjunto de los enteros negativos, ordenados de menor a mayor, o el conjunto $\{\cos x\text{: } 0 \leq x \leq \pi\}$, ordenado por la magnitud del argumento $x$. Si $\kappa$ es un ordinal cualquiera, el conjunto

1 Conviene observar que este distingo es tan ajeno al alemán como al castellano. En la vida diaria, 'Anzahl' se usa en frases como "eine große Anzahl Kinder" ("un gran número de niños"), "eine Anzahl von 100" ('un centenar'). El uso matemático en tiempos de Cantor puede ilustrarse con el siguiente pasaje de Lipschitz, que Kronecker cita con aprobación en "Sobre el concepto de número" (1887, p. 342n.; cursiva mía):

> Wenn man bei der Betrachtung getrennter Dinge von der Merkmalen absieht, durch welche sich die Dinge unterscheiden, so bleibt der Begriff der *Anzahl* der betrachteten Dinge zurück.

> Cuando en la consideración de cosas separadas se prescinde de los caracteres que las distinguen resta el concepto del *número* de las cosas consideradas.

'Anzahl' se caracteriza aquí prácticamente en los mismos términos en que Cantor (1895/97, § 1; GA, p. 282) define 'Kardinalzahl' ('número cardinal'). Por esos mismos años, Frege confirió a 'Anzahl' un significado técnico especial distinto del que le da Cantor, mediante esta definición: "La *Anzahl* correspondiente al concepto F es la extensión del concepto 'equinumeroso con el concepto F'" (1884, § 68; pp. 79–80). También esta noción es afín a la cantoriana de 'número cardinal'.

2 El orden lineal así definido concuerda con lo que Cantor llama un "orden simple" (GA, pp. 296, 444). Sus definiciones tardías de conjunto bien ordenado presuponen explícitamente que se trata de un conjunto ordenado simplemente (GA, pp. 312, 444).

[κ] es lo que Cantor llama un *conjunto bien ordenado* por la relación <.[3] En el referido escrito Nº 5 Cantor explica que un conjunto bien ordenado (*eine wohlgeordnete Menge*) es un conjunto bien definido cuyos elementos están ordenados linealmente de tal modo que:

**BO1** Hay un primer elemento, esto es, un elemento que precede a todos los otros.

**BO2** Todo elemento que precede a otros tiene un sucesor inmediato, esto es, un elemento que le sigue y precede a cualquier otro elemento precedido por él.

**BO3** Si *A* es una parte no vacía —finita o infinita— del conjunto y los elementos del conjunto que siguen a todos los elementos de *A* forman otra parte no vacía *B*, entonces *B* tiene un primer elemento (esto es, hay un elemento de *B* que precede a cualquier otro elemento de *B*).

Esta definición de conjunto bien ordenado (GA, p. 168) equivale a la definición más sencilla que Cantor propuso en una carta a Dedekind el 28 de julio de 1899 (GA, p. 444) y que figura en los libros de texto actuales:[4]

3 Como luego se verá, las condiciones **BO1–3** con que Cantor inicialmente define el concepto de buen orden están calculadas para garantizar que el mismo sea aplicable a la sucesión de ordinales resultante de los "principios generadores" enunciados en Cantor 1883 (*vide infra*, pp. 36 s.). Por eso, quizás, Cantor no se molesta en demostrar que la sucesión de los ordinales —y, por ende, cada segmento inicial suyo— está bien ordenada por la relación <. Para los lectores a quienes esto no les parezca obvio, Zermelo prueba —en una nota a su edición de los escritos de Cantor— que la sucesión de los ordinales cumple la condición **BO** (Cantor, GA, p. 208, n.4). Sea *K* cualquier conjunto de ordinales. Sea $K^* = \{\xi: \forall\zeta(\zeta \in \mathrm{K} \rightarrow \xi < \zeta)\}$ ($K^*$ es el conjunto de los ordinales menores que cada elemento de *K*). Entonces, o bien *K* contiene el primer ordinal y éste es su primer elemento; o bien no lo contiene, y entonces $K^* \neq \varnothing$. En este caso, o bien $K^*$ tiene un elemento máximo $\mu$, y el ordinal $\mu + 1$, determinado por el primer principio generador, es el primer elemento de *K*; o bien $K^*$ no tiene un elemento máximo, en cuyo caso el segundo principio generador determina un ordinal límite, sucesor próximo de todos los elementos de $K^*$, que es el primer elemento de *K*. Así, en las tres alternativas posibles, *K* tiene un primer elemento.

4 Demuestro la equivalencia de ambas definiciones en el Apéndice I. El lector habrá advertido que la condición **BO1** presupone que el conjunto en cuestión no está vacío. Por otra parte, ∅ cumple trivialmente la condición **BO**: como ∅ *no tiene* ninguna parte

**BO** Un conjunto linealmente ordenado $M$ se dice *bien ordenado* si cada parte no vacía $A \subseteq M$ tiene un primer elemento.

Como es obvio, un conjunto infinito numerable, esto es, un conjunto equinumeroso con el conjunto [ω] de los enteros positivos finitos, queda bien ordenado por cualquier aplicación biyectiva de [ω] sobre él. Pero ello, evidentemente, no implica que cualquier conjunto admita un ordenamiento de este tipo. Por ejemplo, como el lector fácilmente adivinará, no se conoce ningún procedimiento para bien ordenar el conjunto infinito de los puntos de un cuadrado. Ello no obstante, a la citada explicación del concepto de buen orden por Cantor sigue la siguiente extraordinaria declaración:

> El concepto de *conjunto bien ordenado* resulta ser fundamental para toda la teoría de las variedades (*Mannigfaltigkeitslehre*). Que siempre es posible reducir cada conjunto *bien definido* a la *forma* de un conjunto *bien ordenado* es una ley del pensamiento, a mi modo de ver, básica y fecunda, y especialmente notable por su universalidad, a la cual retornaré en un trabajo posterior.
>
> (Cantor, GA, p. 169).

Más tarde, Cantor no volverá a calificar esta proposición tan poco evidente como "ley del pensamiento", y hasta hará un intento por demostrarla, que explico en el Capitulo 1.6. En el Capítulo 1.7 y en el Apéndice VI consideraré las demostraciones publicadas por Zermelo en 1904 y 1908. Desde entonces la proposición se conoce como el Teorema del Buen Orden y así me referiré a ella en lo sucesivo. Veremos que, al igual que el célebre Postulado V de Euclides, no puede demostrársela a menos que se acepte otra proposición que no es más obvia. Me parece, por eso, que Cantor puso de manifiesto un certero instinto matemático cuando intentó hacer pasar el Teorema del Buen Orden por un principio del pensamiento que no requiere demostración (así como Euclides mostró su genio al dar rango axiomático al Postulado V). Para su programa, el Teorema del Buen Orden era indispensable: la sucesión de los ordinales alcanza para enumerar todo lo que se presente en

---

no vacía, no tiene una que carezca de un primer elemento. **BO** equivale a **BO1** $\wedge$ **BO2** $\wedge$ **BO3** si y sólo si $M \neq \varnothing$. Como Cantor no contemplaba la existencia de $\varnothing$ entendía seguramente que el conjunto al que se refiere la condición **BO** no estaba vacío.

la naturaleza corpórea y espiritual si —pero sólo si— cada conjunto puede ordenarse bien. A menos que todo conjunto sea bien ordenable no se justifica el tono triunfalista del siguiente pasaje:

> Una de las tareas más importantes de la teoría de los conjuntos, que creo haber resuelto en lo principal en [el escrito Nº 5 de 1883], consiste en la exigencia de determinar las distintas valencias o potencias [esto es, las numerosidades] de las variedades presentes en la totalidad de la naturaleza, en la medida en que ésta se abre a nuestro conocimiento. Lo he logrado mediante la formación del *concepto* general *del enumerador de un conjunto bien ordenado*, o, lo que es lo mismo, del *concepto de número ordinal.*
>
> (Cantor 1887; GA, pp. 387s.)

En el citado escrito Nº 5 Cantor formula tres principios que según él presiden la generación de los ordinales. Observa que la serie de los enteros positivos finitos 1, 2, 3, . . . , *n*, . . . surge del repetido "postular y reunir" (*Setzung und Vereinigung*) unidades que se consideran iguales. "El número *n* es a la vez la expresión de un determinado número finito de tales postulaciones sucesivas y de la reunión de las unidades postuladas en un todo" (GA, p. 195). La doble operación descrita ilustra un primer principio generador (*Erzeugungsprinzip*): "el principio de la adición de una unidad a un número ya formado", el cual, según Cantor, desempeña también un papel esencial en la generación de los enteros transfinitos. Los números formados de este modo constituyen lo que Cantor llama la *primera clase de números* o la Clase (I).

> Aunque sería contradictorio hablar de un número máximo de la Clase (I), no hay nada chocante (*nichts Anstößiges*) en concebir un *nuevo* número, que llamaremos ω, y que ha de ser la expresión de que todo el agregado (I) está dado en su sucesión natural conforme a la ley (al modo como *n* es la expresión de que un cierto número finito de unidades está reunido en un todo).
>
> (Cantor, GA, p. 195)

Si a la postulación del número ω siguen nuevas postulaciones de la unidad se obtienen, mediante el primer principio, los nuevos números ω + 1, ω + 2, . . . , ω + *n*, . . . No habrá entre ellos un número máximo que siga a todos los otros, pero se puede pensar en un número nuevo 2ω, que exprese

la reunión de aquéllos en un todo. Aplicando el primer principio a $2\omega$ se obtienen los números $2\omega + 1$, $2\omega + 2, \ldots,$ $2\omega + n, \ldots$ [5] Al lector desprevenido esta proliferación de números le parecerá un truco de ilusionista, pero vimos arriba, a propósito de la formación de los conjuntos derivados de ciertos conjuntos de puntos, que la sucesión descrita aquí en abstracto por Cantor admite aplicaciones perfectamente bien definidas. (Piénsese también en esto: $\omega$ es el enumerador del conjunto de los números pares; si a ese conjunto vamos agregando los impares $1, 3, 5, \ldots$ y prescribimos que se los enumere "en orden de llegada", es claro que se necesitarán los enumeradores $\omega + 1$, $\omega + 2$, $\omega + 3, \ldots$ ).

> La función lógica que nos ha suministrado los dos números $\omega$ y $2\omega$ es manifiestamente distinta de la del *primer* principio generador. La llamo el *segundo principio generador* de números reales enteros, que defino con más precisión así: cuando hay una sucesión determinada cualquiera de números enteros reales definidos, ninguno de los cuales es el mayor de todos ellos, se crea en virtud de este segundo principio generador un número nuevo, que se concibe como *límite* de todos esos números, vale decir, se lo define como el próximo sucesor de todos ellos (*als die ihnen allen nächst größere Zahl*).
>
> (Cantor, GA, p. 196)

Corrientemente se llama *ordinal sucesor* a uno generado conforme al *primer* principio y *ordinal límite* a uno generado conforme al *segundo*. Un ordinal límite no es el sucesor *inmediato* de otro ordinal; en cambio, cabe describirlo como el sucesor *próximo* de los ordinales que le preceden, en cuanto él mismo precede a cualquier otro ordinal mayor que todos sus predecesores. Cantor observa que el ordinal $\omega$ y los primeros ordinales sucesores y ordinales límites que le siguen tienen la propiedad de que el conjunto de sus respectivos predecesores es numerable.[6] Cantor hace de esta propiedad la característica esencial de lo que llama la *segunda clase de números* o Clase

5 Como luego veremos, la multiplicación de ordinales no es conmutativa si uno de los factores es transfinito. En particular, $2\omega = \omega \neq \omega 2$. Por eso, Cantor llamará más tarde $\omega 2$ al número introducido aquí bajo la designación '$2\omega$' (GA, p. 389).

6 Desde luego, tienen esta propiedad los ordinales transfinitos $\omega 2$ y $\omega^2$ que he aducido como ejemplos, puesto que ellos enumeran, respectivamente, (i) el conjunto ordenado

(II). Dicha propiedad puede verse como un *tercer* principio "de inhibición o limitación" (*Hemmungs- oder Beschränkungsprinzip*), en virtud del cual:

($a_1$) Todo ordinal transfinito $\kappa$ formado conforme a los principios generadores tiene un conjunto de predecesores $[\kappa]$ equinumeroso con $[\omega]$.

La Clase (II) está formada por $\omega$ y los ordinales sucesores y ordinales límites que cumplen con la exigencia ($a_1$). Cumple además las condiciones siguientes:

($b_1$) El agregado de todos los números de la Clase (II) es más numeroso que el conjunto $[\omega]$ de los números de la Clase (I); y

($c_1$) no existe una numerosidad intermedia mayor que la de la Clase (I) y menor que la de la Clase (II).[7]

Sea $\omega_1$ el enumerador del conjunto de los ordinales de las Clases (I) y (II), ordenados de menor a mayor. Consíderese ahora la condición

($a_2$) $[\kappa]$ es equinumeroso con $[\omega_1]$.

Los ordinales generados conforme a los principios cantorianos bajo esta condición forman una Clase (III) tal que

($b_2$) el agregado de todos los ordinales de la Clase (III) es más numeroso que el conjunto $[\omega_1]$ de los números de la Clase (II); y

($c_2$) no hay una numerosidad intermedia mayor que la de la Clase (II) y menor que la de la Clase (III).

En forma análoga, se forman las Clases (IV), (V), . . . , de ordinales, tales

---

que forman todos los impares 1, 3, 5, . . . seguidos por todos los pares 2, 4, 6, . . . y (ii) el conjunto ordenado que forman los números primos 2, 3, 5, 7, 11, . . . seguidos por sus respectivos cuadrados, cubos, . . . ; y, evidentemente, cada uno de estos conjuntos puede biyectarse sobre el conjunto de los predecesores de $\omega$ (con el cual se confunde, si olvidamos el orden).

7 En el Apéndice III bosquejo su demostración de estas dos proposiciones.

que, para cada entero positivo *n*, (i) el conjunto [$\xi$] de los predecesores de cualquier ordinal $\xi$ de la clase (*n*+1)-ésima clase es equinumeroso con el conjunto de los ordinales de la clase *n*-ésima; (ii) la clase *n*-ésima es menos numerosa que la clase (*n*+1)-ésima; y (iii) no existe una numerosidad intermedia entre las de las clases *n*-ésima y la (*n*+1)-ésima. Por otra parte, como Zermelo observa en su edición de los escritos de Cantor (GA, p. 199), los principios citados no bastan para establecer la existencia de una clase $\omega$-ésima de ordinales.

En sus obras más maduras, Cantor se refiere a esta serie ascendente de numerosidades infinitas conjurada por los sucesores de $\omega$ como a una serie numérica transfinita distinta de la de los ordinales, aunque asociada a ella: la sucesión de los números cardinales transfinitos o alephs (llamados así por la letra hebrea aleph, $\aleph$, que Cantor adoptó para nombrarlos). Para Cantor, el cardinal de un conjunto epitomiza lo que él llamaba su potencia o valencia y aquí he llamado su *numerosidad*. Si olvidamos todas las particularidades de los elementos de un conjunto *K*, recordando únicamente que son diferentes entre sí, obtenemos, según Cantor, un concepto que se aplica igualmente a *K* y a cualquier otro conjunto equinumeroso con *K*. Este concepto es lo que él llama el número cardinal (*Kardinalzahl*) de *K*. Si *K* es menos numeroso que otro conjunto *L*, decimos que el cardinal de *K* es menor que el cardinal de *L*. El conjunto de los cardinales finitos está bien ordenado por la relación '*x* es menor que *y*' —que simbolizaremos: $x \prec y$—, puesto que, evidentemente, hay un isomorfismo[†] de órdenes entre los cardinales finitos, ordenados por la relación $\prec$, y los enumeradores de los conjuntos respectivos, ordenados por la relación <. (Usaré también la abreviatura '$x \preccurlyeq y$' para decir '$x \prec y$ o $x = y$'.) En virtud de este isomorfismo los cardinales finitos se identifican naturalmente con los respectivos ordinales y se conocen por los mismos nombres que éstos. Pero, como hemos visto, tal correspondencia no existe en el dominio transfinito. Cantor insiste, por ello, en que los cardinales transfinitos son otra clase de objetos que los ordinales y los designa con nombres peculiares. Así, el cardinal del conjunto [$\omega$] —o del conjunto de los predecesores de cualquier ordinal de la segunda clase— se llama $\aleph_0$ (aleph cero); el cardinal del conjunto de los ordinales de segunda clase —o del conjunto de los predecesores de cualquier ordinal de la tercera clase— se llama $\aleph_1$, etc.

Como dije arriba, Cantor demostró que entre $\aleph_0$ y $\aleph_1$ no puede haber

un cardinal intermedio (véase el Apéndice II). Esta proposición se deja generalizar a cualquier ordinal $\kappa \geq 1$: en otras palabras, si $\aleph_\kappa$ y $\aleph_{\kappa+1}$ son los cardinales de los conjuntos $K$ y $K'$, respectivamente, no puede haber un conjunto $K''$ más numeroso que $K$ y menos numeroso que $K'$. Sea $K$ un conjunto infinito cualquiera. Si $K$ admite un buen orden hay un ordinal que lo enumera, y $K$ es equinumeroso con el conjunto de los predecesores de ese ordinal. Por lo tanto, el cardinal de $K$ es $\aleph_\lambda$ (para algún ordinal $\lambda$). Pero si $K$ no admite un buen orden, no hay ningún ordinal que lo enumere y el cardinal de $K$ no pertenece a la secuencia de los alephs. Peor aún, como demostrará Hartogs (1915), a menos que todo conjunto sea bien ordenable, habrá numerosidades incomparables, esto es, pares de conjuntos $K$ y $K'$ tales que $K$ no es ni más ni menos numeroso que $K'$ ni equinumeroso con él.

Está claro, entonces, que la pretensión de Cantor de haber provisto un sistema numérico suficiente para medir todas las numerosidades de la naturaleza depende enteramente del Teorema del Buen Orden. Por otra parte, si —como él— lo damos por supuesto, obtenemos sólo la seguridad de que todas esas numerosidades son medibles a ojos de Dios, puesto que nosotros no somos capaces de bien ordenar ni siquiera el conjunto de los puntos de un intervalo, ni, en general, ningún conjunto que no sea finito o numerable. Por esta razón, la clara jerarquía de los alephs correspondientes a las sucesivas clases de ordinales transfinitos resulta de poco interés fuera del recinto cerrado de la teoría cantoriana si no se la vincula a la serie ascendente de numerosidades generada en el primer frente del avance al transfinito por la operación $K \mapsto \mathcal{P}K$. Como veremos un poco más adelante, es fácil probar que el conjunto $\mathbb{R}$ de los números reales es equinumeroso con $\mathcal{P}[\omega]$. La Hipótesis del Continuo significa, entonces, que no hay conjuntos de numerosidad intermedia entre la de $[\omega]$ y la de $\mathcal{P}[\omega]$, de modo que el cardinal de $\mathcal{P}[\omega]$ —y de $\mathbb{R}$— es precisamente $\aleph_1$. La Hipótesis del Continuo constituye así el primer eslabón de una cadena que vincularía la jerarquía de los alephs a las numerosidades generadas por la operación $K \mapsto \mathcal{P}K$. Por eso, Cantor se empeñará mucho en demostrar la Hipótesis del Continuo y se sentirá atribulado por su fracaso. La cadena se completa con la llamada Hipótesis Generalizada del Continuo, formulada en el espíritu de Cantor por sus sucesores:

**HGC** Si $\aleph_\kappa$ es el cardinal de un conjunto $K$, el cardinal de $\mathcal{P}K$ es $\aleph_{\kappa+1}$, cualquiera que sea el ordinal $\kappa$.

**HGC** implica el Teorema del Buen Orden (pero no es implicada por él). Por lo tanto, sólo cabría admitirla como un principio que no se demuestra. Volveremos sobre esto en la Tercera Parte. Es claro que bajo **HGC** los alephs darían abasto para cubrir todas las numerosidades de la "naturaleza corpórea y espiritual".

Con clara conciencia de lo que se espera de un "número" de veras digno de ese nombre, Cantor pone especial cuidado en definir las operaciones aritméticas elementales aplicables a sus ordinales y cardinales. No podemos entrar en detalles, pero un par de indicaciones darán una idea de este asunto.

Aunque la aritmética transfinita ordinal precede históricamente a la cardinal, es más fácil explicarla si hablamos primero de ésta. Este es el orden que Cantor mismo sigue en su presentación final del tema (Cantor 1895/97; GA, pp. 282ff.), en la que me basaré. Como ya he dicho, para Cantor el cardinal de un conjunto $K$ es el concepto abstracto que se forma olvidando todas las propiedades y relaciones de los elementos de $K$ excepto una sola, a saber, que cada uno de esos elementos es idéntico consigo mismo y diferente de todos los demás. Como es obvio, el mismo concepto se obtiene a partir de cualquier conjunto equinumeroso con $K$.[8] Para referirme a los cardinales —finitos o transfinitos— usaré como Cantor letras góticas, $\mathfrak{a}$, $\mathfrak{b}$, $\mathfrak{c}$, . . . Cantor designa el cardinal del conjunto $K$ mediante el símbolo $\overline{\overline{K}}$, pero aquí lo llamaremos $|K|$, como es ahora habitual. Por definición, la suma $\mathfrak{a} + \mathfrak{b}$ de dos cardinales $\mathfrak{a}$ y $\mathfrak{b}$ es el cardinal de la unión de dos conjuntos disjuntos cualesquiera tales que $\mathfrak{a}$ y $\mathfrak{b}$ sean sus repectivos cardinales. En otras palabras, si $\mathfrak{a} = |A|$ y $\mathfrak{b} = |B|$ y $A \cap B = \varnothing$, $\mathfrak{a} + \mathfrak{b} = |A \cup B|$. El producto $\mathfrak{ab}$ de dos cardinales $\mathfrak{a}$ y $\mathfrak{b}$ se define así: si $\mathfrak{a} = |A|$ y $\mathfrak{b} = |B|$, sea $A \times B$ el conjunto de todos los pares ordenados $\langle a,b \rangle$, tales que $a \in A$ y $b \in B$; entonces $\mathfrak{ab} = |A \times B|$. Cantor muestra que bajo las condiciones antedichas, $\mathfrak{ab}$ es asimismo el cardinal del conjunto construido formando la unión de $\mathfrak{a}$ conjuntos disjuntos

8 La caracterización cantoriana de los cardinales recuerda, sin duda, la definición de Frege, según la cual el cardinal de una clase $C$ es la clase de todas las clases equinumerosas con $C$, pero no debe confundirse con ella. La definición de Frege identifica el cardinal de la clase $C$ con la extensión del concepto con el cual lo identificaría Cantor si $C$ fuese en efecto un conjunto.

equinumerosos con $B$; en otras palabras, si $\mathfrak{a} = |A|$, y para todo $h, k \in A$, $\mathfrak{b} = |B_h|$ y $B_h \cap B_k = \varnothing$ si $h \neq k$, entonces $\mathfrak{ab} = |\bigcup_{k\in A} B_k|$.[9] Es fácil verificar que la adición y la multiplicación así definidas son operaciones conmutativas y asociativas: $\mathfrak{a} + \mathfrak{b} = \mathfrak{b} + \mathfrak{a}$; $\mathfrak{a} + (\mathfrak{b} + \mathfrak{c}) = (\mathfrak{a} + \mathfrak{b}) + \mathfrak{c}$; $\mathfrak{ab} = \mathfrak{bc}$; $\mathfrak{a}(\mathfrak{bc}) = (\mathfrak{ab})\mathfrak{c}$; y que la multiplicación es distributiva sobre la adición: $\mathfrak{a}(\mathfrak{b} + \mathfrak{c}) = \mathfrak{ab} + \mathfrak{ac}$. Asimismo, si llamamos 0 al cardinal de $\varnothing$ y 1 al de un conjunto unitario, es claro que $\mathfrak{a} + 0 = \mathfrak{a}$, $0\mathfrak{a} = 0$ y $1\mathfrak{a} = \mathfrak{a}$. Cantor mostró además que si $\mathfrak{a} \preccurlyeq \mathfrak{b}$, hay un cardinal $\mathfrak{c}$ tal que $\mathfrak{b} = \mathfrak{a} + \mathfrak{c}$, y que si $\mathfrak{a} \preccurlyeq \mathfrak{b}$ y $\mathfrak{c} \preccurlyeq \mathfrak{d}$, entonces $\mathfrak{a} + \mathfrak{c} \preccurlyeq \mathfrak{b} + \mathfrak{d}$ y $\mathfrak{ac} \preccurlyeq \mathfrak{bd}$. Hasta aquí, pues, la aritmética transfinita cardinal no difiere de la finita. Por otra parte, se puede demostrar que si a es un cardinal transfinito y $n$ es un cardinal finito, $\mathfrak{a} + n = \mathfrak{a}$. Hessenberg (1906) estableció que, para cualquier ordinal $\kappa$, $\aleph_\kappa + \aleph_\kappa = \aleph_\kappa$, de lo cual se deduce fácilmente que si $\kappa < \lambda$, $\aleph_\kappa + \aleph_\lambda = \aleph_\lambda$.[10]

Para definir la exponenciación de cardinales Cantor define lo que llama ‘cubrimiento’ de un conjunto $N$ con un conjunto $M$ (*Belegung von N mit M*), como “una ley en virtud de la cual se le asocia a cada elemento $n$ de $N$ un determinado elemento de $M$, pudiendo usarse repetidas veces un mismo elemento de $M$.” Un cubrimiento de $N$ con $M$ es pues precisamente lo que hoy se llama una *aplicación*[†] de $N$ en $M$, y usaré este término en lugar del utilizado por Cantor. Por definición, si $\mathfrak{a} = |A|$ y $\mathfrak{b} = |B|$, $\mathfrak{a}^{\mathfrak{b}}$ es el cardinal del conjunto de todas las aplicaciones de $B$ en $A$. Cantor muestra que la operación así definida obedece a las leyes familiares de la exponenciación finita: $\mathfrak{a}^{\mathfrak{b}}\mathfrak{a}^{\mathfrak{c}} = \mathfrak{a}^{\mathfrak{b}+\mathfrak{c}}$, $(\mathfrak{a}^{\mathfrak{b}})^{\mathfrak{c}} = a^{\mathfrak{bc}}$.

Es claro que, si $\mathfrak{f}$ es finito, $\mathfrak{f} \prec 2^{\mathfrak{f}}$. Cantor demostró que $\mathfrak{f} \prec 2^{\mathfrak{f}}$ también si $\mathfrak{f}$ es infinito. Por su importancia, este aserto se llama comúnmente el *Teorema de Cantor*. La siguiente prueba, sencilla y elegante, se publicó en Cantor 1890/91 (GA, pp. 279s.). Sea $K$ un conjunto infinito y $F$ el conjunto de todas las aplicaciones de $K$ en $\{0,1\}$. Sea $f{:}K \to F$ una aplicación inyectiva. Para mayor claridad escribo $f_x$ en vez de $f(x)$, para designar el valor asignado por $f$ a $x \in K$. Considérese la aplicación $\varphi$ de $K$ en $\{0,1\}$ definida así: $\varphi(x) = 0$ si $f_x(x) = 1$ y $\varphi(x) = 1$ si $f_x(x) = 0$. Obviamente, $\varphi$ difiere de $f_x$ para cada $x \in K$. Por lo tanto, $\varphi$ es un elemento de $F$ que cae fuera del alcance

9 En estas explicaciones he empleado la terminología y simbolismo actuales.

10 Estos resultados se demuestran en todos los buenos manuales de teoría de conjuntos, por ejemplo, en Levy 1979, III.3.

de la aplicación inyectiva $f$. Por ende, $f$ no es una biyección de $K$ en $F$. Como este argumento es aplicable a cualquier inyección de $K$ en $F$, $|K| \prec |F| = 2^{|K|}$. El argumento ilustra un modo de razonar —el llamado "método diagonal"— que Russell utilizará en la construcción de la paradoja que lleva su nombre (Capítulo 1.6) y Gödel en la demostración de su primer teorema de incompletud (Capítulo 2.10).[11]

Ahora mostraré que si $K$ es un conjunto cualquiera, el conjunto $\mathcal{P}K$ de todas las partes de $K$ es más numeroso que $K$. Este resultado asegura una provisión inagotable de cardinales transfinitos diferentes (si suponemos, como Cantor al parecer suponía, que existe un conjunto infinito y que la existencia de un conjunto implica la existencia del conjunto de sus partes). Es una consecuencia inmediata del Teorema de Cantor. En efecto, cada elemento $H$ de $\mathcal{P}K$ —es decir, cada conjunto $H \subseteq K$— está asociado a una determinada aplicación de $K$ en $\{0,1\}$, la *función característica* $\chi_H$ definida así para cada $k \in K$: $\chi_H(k) = 1$ si $k \in H$ y $\chi_H(k) = 0$ si $k \notin H$.[12] Por otra parte, cada aplicación $\chi{:}K \to \{0,1\}$ determina un elemento de $\mathcal{P}K$, a saber,

11 Aunque suele atribuírsele a Cantor, el método diagonal fue empleado antes por Paul du Bois-Reymond (1875, p. 365n.) para demostrar que, dada una secuencia de funciones reales $\lambda_1, \lambda_2, \dots$ que crecen monotónicamente sin cota ($\lim_{x\to\infty} \lambda_k(x) = \infty$), pero cada cual más lentamente que la anterior ($\lim_{x\to\infty} \frac{\lambda_{k+1}(x)}{\lambda_k(x)} = 0$), existe siempre una función real $\lambda$ que crece sin cota, pero más lentamente que cualquiera de las $\lambda_k$. En términos generales, el método puede describirse así: dado un conjunto $\mathcal{D}$ y una "función de dos variables" $g$ definida en $\mathcal{D}^2$, pasamos a considerar la *diagonalización* de $g$, esto es, la "función de una variable" $h$ definida en $\mathcal{D}$ por la condición $h(x) = g(x,x)$. El nombre 'diagonalización' se justifica fácilmente en caso que $\mathcal{D}$ sea finito o enumerable, ya que entonces los valores de $g$ se pueden desplegar naturalmente en una matriz cuya diagonal exhibe los valores de $h$. La popular demostración de la indenumerabilidad del continuo publicada en Cantor 1890/91 (GA, pp. 278f.) ilustra el método diagonal en esta forma: Sea $\varphi$ una aplicación inyectiva de $[\omega]$ en $\mathbf{I} = (0,1]$. Sea $0{,}a_{k0}a_{k1}a_{k2}\dots a_{kn}\dots$ la representación de $\varphi(k)$ mediante su expansión decimal infinita. Sea, para todo ordinal finito $h$, $a_h = 1$ si $a_{hh} \neq 1$ y $a_h = 2$ si $a_{hh} = 1$. Entonces, $0{,}a_0a_1a_2\dots a_n\dots$ es la expansión decimal infinita de un elemento de $\mathbf{I}$ que cae fuera del alcance de la aplicación $\varphi$, puesto que difiere de $\varphi(k)$ en el $k$-ésimo decimal. Por lo tanto, no existe una aplicación biyectiva de $[\omega]$ en $\mathbf{I}$, de suerte que $[\omega]$ es menos numeroso que $\mathbf{I}$.

12 Mas adelante, cuando hablemos de Gödel, será preferible definir con él la función característica $\chi_H$ de un conjunto $H$ por la condición $\chi_H(k) = 0$ si $k \in H$ y $\chi_H(k) = 1$ si $k \notin H$. Como es obvio, cualquier decisión en esta materia es puramente convencional.

el conjunto $\{k \in K{:}\,\chi(k) = 1\}$. Por lo tanto, $|\mathcal{P}K| = 2^{|K|} > |K|$. En particular, el cardinal del conjunto de las partes del conjunto $[\omega]$ de los ordinales finitos es $2^{\aleph_0}$, puesto que $\aleph_0$ es, por definición, el cardinal de $[\omega]$. $2^{\aleph_0}$ es también el cardinal del continuo, como paso a demostrar. Sea $\mathbf{I} = (0,1] \subseteq \mathbb{R}$ ($\mathbf{I}$ es el conjunto de los números reales $x$ tales que $0 < x \leq 1$). Vimos en el Capitulo 1.3 que $\mathbf{I}$ es equinumeroso con un continuo de cualquier dimensión. Sea $a \in \mathbf{I}$. Entonces $a$ puede representarse mediante su expansión decimal infinita $0{,}a_0a_1a_2 \ldots a_n \ldots$, donde para cada ordinal finito $k$, $a_k$ es uno de los diez dígitos y hay un entero $h > k$ tal que $a_h \neq 0$. También puede representárselo mediante su expansión binaria infinita, cuya definición es la misma que la anterior, excepto que para cada ordinal finito $k$, $a_k$ es uno de los dos dígitos del sistema binario: 0 ó 1. Ahora bien, una tal expansión binaria infinita puede leerse como la lista de los valores de la aplicación de $[\omega]$ en $\{0,1\}$ por $k \mapsto a_k$, que es la función característica del conjunto $H$ de ordinales finitos definido por la condición: $k \in H$ si $a_k = 1$ y $k \notin H$ si $a_k = 0$. La correspondencia así establecida entre las referidas expansiones binarias y ciertos subconjuntos de $[\omega]$ determina una aplicación inyectiva $f{:}\, \mathbf{I} \rightarrow \mathcal{P}[\omega]$. $f$ no es biyectiva, pues su alcance comprende solamente los subconjuntos *infinitos* de $[\omega]$. En efecto, si $H \subseteq [\omega]$ es finito, $H$ contiene un elemento máximo $n$, de modo que $\chi_{\mathrm{H}}(m) = 0$ para todo ordinal finito $m > n$ y la lista de valores de $\chi_{\mathrm{H}}$ no es la expansión binaria infinita de ningún número. Sin embargo, como hay solamente un número finito de partes diversas de $[\omega]$ que tienen el mismo elemento máximo $n$ es claro que el conjunto de las partes finitas de $[\omega]$ es numerable. Por lo tanto, si U es la unión del intervalo $(0,1]$ con un conjunto numerable, por ejemplo, con el conjunto $\{x{:}\ x$ es racional y $1 < x < 2\}$, se puede definir una aplicación biyectiva $f'{:}\ \mathrm{U} \rightarrow \mathcal{P}[\omega]$, tal que $f$ es la restricción de $f'$ a $(0,1]$. Por lo tanto, el cardinal de U es idéntico al cardinal de $\mathcal{P}[\omega]$. Como $(0,1]$ y $(0,2]$ son equinumerosos y $(0,1] \subseteq \mathrm{U} \subseteq (0,2]$, es claro que $|(0,1]| = |\mathrm{U}| = |\mathcal{P}[\omega]| = 2^{\aleph_0}$. En vista de esto, la Hipótesis del Continuo puede enunciarse —como es habitual— así:

$$\aleph_1 = 2^{\aleph_0} \qquad \textbf{(HC)}$$

El argumento anterior ofrece una prueba más de que el conjunto de los ordinales finitos es menos numeroso que el continuo.

Para entender la aritmética ordinal conviene explicar primero el concepto

cantoriano de *tipo de orden*. Sea $K_i$ un conjunto no vacío y $P_i$ una relación binaria entre miembros de $K_i$ ($i$ = 1, 2). Decimos que la estructura $\langle K_1,P_1\rangle$ es isomórfica a la estructura $\langle K_2,P_2\rangle$ si hay entre ellas un isomorfismo, esto es, una aplicación biyectiva $f{:}K_1 \to K_2$ tal que, para todo $x$, $y \in K_1$, $x$ tiene la relación $P_1$ con $y$ si y sólo si $f(x)$ tiene la relación $P_2$ con $f(y)$ (simbólicamente: $xP_1y \Leftrightarrow f(x)P_2f(y)$). Si $P_1$ es una relación de orden lineal, el *tipo de orden* de $\langle K_1,P_1\rangle$ es, según Cantor, el concepto que se forma al hacer abstracción de todas las particularidades de los elementos de $K_1$ excepto (i) la identidad de cada uno consigo mismo y su diversidad de los otros y (ii) el orden entre ellos determinado por la relación $P_1$. Obviamente, este concepto es común a todas las estructuras isomórficas a $\langle K_1,P_1\rangle$.[13] Cantor llama *similares* (*ähnlich*) a los conjuntos que tienen un mismo tipo de orden, término que usaré en lo sucesivo. En 1897 Cantor *define* los ordinales como los tipos de orden de los conjuntos bien ordenados (GA, p. 321). En particular, cada ordinal $\kappa$, formado con arreglo a los tres principios enunciados por Cantor en 1883, habrá de identificarse simplemente con el tipo de orden de los conjuntos bien ordenados similares con $\langle[\kappa],<\rangle$. Volveré sobre esto después de explicar la adición y multiplicación de tipos de orden.

Cantor designa los tipos de orden con minúsculas griegas. La suma $\alpha + \beta$ de dos tipos de orden $\alpha$ y $\beta$ se define como el tipo de orden de la estructura $\langle K,P\rangle$ construida así: (i) tómese una estructura $\langle K_1,P_1\rangle$ de tipo $\alpha$ y una estructura $\langle K_2,P_2\rangle$ de tipo $\beta$, tales que $K_1 \cap K_2 = \varnothing$; (ii) sea $K = K_1 \cup K_2$; (iii) defínase así el orden lineal $P$ en $K$: $xPy$ si y sólo si $x$, $y \in K_i$ y $xP_iy$, o si $x \in K_1$ e $y \in K_2$. El producto $\alpha\beta$ de $\alpha$ y $\beta$ se define como el tipo de orden de la estructura $\langle K,P\rangle$ construida así: (i) tómese una estructura $\langle K_1,P_1\rangle$ de tipo $\alpha$ y una estructura $\langle K_2,P_2\rangle$ de tipo $\beta$, tales que $K_1 \cap K_2 = \varnothing$; (ii) fórmese una colección de copias disjuntas de $\langle K_1,P_1\rangle$, una para cada elemento de $K_2$; sea en particular $\langle K_1(x),P_1\rangle$ la copia correspondiente a $x \in K_2$; (iii) sea $K$ la unión de las copias de $K_1$ ($K = |\bigcup_{x\in K1}K_1(x)|$); (iv) defínase así el orden lineal $P$ en $K$: $uPw$ si y sólo si $u$, $w \in K_1(x)$ para un cierto $x \in K_2$, o si $u \in K_1(x)$ y $w \in K_1(y)$ y $xP_1y$. Las operaciones así definidas

13 Siguiendo a Cantor, he restringido la definición de tipo de orden a los órdenes lineales; pero puede extendérsela sin más a cualquier estructura que posea algo que pudiéramos llamar un orden. Por ejemplo, es natural decir que los puntos de una circunferencia exhiben un tipo de orden (cíclico) que tienen también los puntos de una recta en el plano proyectivo o los rayos que salen de un punto en el plano euclidiano.

son obviamente asociativas, pero no son conmutativas, como muestran los siguientes ejemplos: (i) $1 + \omega \neq \omega + 1$, puesto que $\omega + 1$ es el tipo de orden de la estructura $\langle[\omega]\cup\{\omega\},<\rangle$ y $1 + \omega$ es el tipo de orden de la estructura $\langle\{1\} \cup ([\omega]\setminus\{1\}),<\rangle = \langle[\omega],<\rangle$ (de modo que $1 + \omega = \omega$). (ii) $2\omega \neq \omega 2$, puesto que $\omega 2$ es el tipo de orden de la estructura $\langle\{\alpha\text{: } \alpha < \omega \text{ y } \alpha \text{ es par}\}\cup\{\beta\text{: } \beta < \omega \text{ y } \beta \text{ es impar}\},P\rangle$, donde $\xi\ P\eta$ si $\xi$ es par y $\eta$ es impar, o si $\xi$ y $\eta$ son pares y $\xi < \eta$, o si $\xi$ y $\eta$ son impares y $\xi < \eta$, de suerte que $\omega 2 \neq \omega$; pero $2\omega$ es el tipo de orden de la estructura $\langle\bigcup_{k<\omega}\{2k,2k+1\},<\rangle$, así que $2\omega = \omega$.[14]

Al definir 'ordinal' como 'tipo de orden de un conjunto bien ordenado' Cantor se desentiende de los principios generadores enunciados en 1883. El nuevo concepto de ordinal es más amplio que el anterior, puesto que cubre también el tipo de orden de un conjunto bien ordenado *K* cuyo cardinal sea $\aleph_\kappa$ para algún $\kappa \geq \omega$, y los principios de 1883 no bastan para generar un ordinal con $\aleph_\omega$ predecesores. Es importante cerciorarse de que ésta es una ampliación natural, que preserva los caracteres estructurales del concepto primitivo. Para facilitar el examen de este asunto, llamaré, por ahora, 'P-ordinales' a los miembros de la sucesión generada en virtud de los principios de 1883 y 'T-ordinales' a los tipos de orden de los conjuntos bien ordenados. Usaré las minúsculas griegas para referirme a ordinales de ambas clases, pero, para evitar confusiones, la letra irá seguida de un asterisco cuando designe un P-ordinal. Veremos que cada P-ordinal puede asimilarse plausiblemente a un determinado T-ordinal y que los T-ordinales que no son P-ordinales se ordenan naturalmente en una sucesión que prolonga la sucesión de los P-ordinales.

Si *k* es un elemento de un conjunto bien ordenado $\langle K,P\rangle$, designo con el símbolo $[k]$ el *segmento* (*Abschnitt*) de *K* determinado por *k*, esto es, el

14 Cantor (1883) ofrece este curioso comentario acerca de la no-conmutatividad de la adición y multiplicación cuando uno de los sumandos o factores es transfinito y el otro no:

> Como se ve claramente, todo depende de la *posición* (*Stellung*) de lo finito hacia lo infinito; si aquél se adelanta, se disuelve en lo infinito y desaparece en su seno; pero si se *modera* y toma su lugar *detrás* de lo infinito, entonces se conserva y se une a éste para formar un infinito nuevo, modificado.
>
> (Cantor, GA, p. 177)

conjunto $\{x \in K: xPk\}$ que forman los predecesores de $k$.[15] Los segmentos de $\langle K,P\rangle$ constituyen una parte de $\mathcal{P}(K)$ bien ordenada por la relación $\subseteq$. Llamémosla $\mathcal{S}(K)$. La aplicación $[k] \mapsto k$ es un isomorfismo entre $\langle \mathcal{S}(K),\subseteq\rangle$ y $\langle K,P\rangle$. Cantor mostró que si $K$ y $H$ son conjuntos bien ordenados por las relaciones $P$ y $Q$, respectivamente, y $\langle K,P\rangle$ no es similar a $\langle H,Q\rangle$, entonces $\langle K,P\rangle$ es similar a un segmento de $\langle H,Q\rangle$, o $\langle H,Q\rangle$ es similar a un segmento de $\langle K,P\rangle$. Este resultado es el Teorema 13N en Cantor 1897.[16] Este teorema permite definir una relación de precedencia $<$ entre los T-ordinales: $\alpha < \beta$ si y sólo si $\alpha$ es el tipo de orden de un segmento de un conjunto bien ordenado cuyo tipo de orden es $\beta$. Es claro, además, que si $\kappa$ es el tipo de orden de un conjunto bien ordenado $\langle K,P\rangle$, cada ordinal $\zeta$ tal que $\zeta < \kappa$ es el tipo de orden de cierto segmento $[z]$ de $\langle K,P\rangle$. La aplicación $\zeta \mapsto [z]$ es un isomorfismo entre $\langle[\kappa],<\rangle$ y $\langle \mathcal{S}(K),\subseteq\rangle$. Por lo tanto, la aplicación $\zeta \mapsto z$ es un isomorfismo entre $\langle[\kappa],<\rangle$ y $\langle K,P\rangle$, de suerte que el T-ordinal $\kappa$ es en efecto el tipo de orden del conjunto de sus predecesores. Supongamos ahora que $\langle K,P\rangle$ es enumerado por el P-ordinal $\kappa^*$. Esto significa, sencillamente que hay un isomorfismo entre $\langle K,P\rangle$ y $\langle[\kappa^*],<\rangle$. En tal caso, $\langle[\kappa^*],<\rangle$ es isomórfico a $\langle[\kappa],<\rangle$. En virtud del teorema 13N, es claro que si $\langle K_1,P_1\rangle$ es un conjunto bien ordenado numerable por un P-ordinal y $\langle K_2,P_2\rangle$ es un conjunto bien ordenado que ningún P-ordinal alcanza para enumerar, $\langle K_1,P_1\rangle$ es similar a un segmento de $\langle K_2,P_2\rangle$, a saber, aquél cuyo tipo de orden corresponde, en virtud del isomorfismo antedicho, al P-ordinal que enumera a $\langle K_1,P_1\rangle$. Estos resultados justifican, sin duda, la identificación de cada P-ordinal $\kappa^*$ con el tipo de orden de los conjuntos bien ordenados que $\kappa^*$ enumera. La sucesión de los T-ordinales que no pueden identificarse así con un P-ordinal obviamente continúa la serie de los que admiten tal identificación. Por lo tanto, de ahora en adelante, hablaré simplemente de 'ordinales' —sin prefijo— y los designaré con minúsculas griegas sin asterisco.

15 Para que $k \in K$ determine un segmento de $K$, Cantor requiere que $k$ no sea el primer elemento de $K$. Esta restricción es innecesaria si se admite la existencia del conjunto vacío. Si $k$ es el primer elemento de $K$, $[k] = \varnothing$; en tal caso, $[k] \subseteq K$, como cualquier otro segmento de $K$.

16 GA, pp. 319s. Bosquejo la demostración en el Apéndice II.

# 1.6 PARADOJAS Y FILOSOFEMAS

En el Capítulo 1.5 reproduje la prueba original del Teorema de Cantor por el método diagonal. Como un corolario del mismo, probé que, si $K$ es un conjunto cualquiera, $|K| < |\mathcal{P}K|$. La prueba del corolario se basó en que cada aplicación de $K$ en $\{0,1\}$ es la función característica $\chi_{\mathrm{H}}$ de un determinado conjunto $H \subseteq K$. Ahora mostraré cómo fundir ambas pruebas en una, de suerte que el aserto de que $|K| < |\mathcal{P}K|$ para todo conjunto $K$ venga a ser sólo otro modo de expresar el Teorema de Cantor, $\mathfrak{f} < 2^{\mathfrak{f}}$. Sea, como antes, $F$ el conjunto de las aplicaciones de $K$ en $\{0,1\}$. Entonces $F = \{\chi_{\mathrm{H}}{:}H \subseteq K\}$ y hay una biyección canónica $g$: $\chi_{\mathrm{H}} \to H$ de $F$ en $\mathcal{P}K$. Si $f{:}x \mapsto f_x$ designa, como en el Capítulo 1.5, una inyección cualquiera de $K$ en $F$, la aplicación compuesta $gf$ ($= g \circ f$) es una inyección cualquiera de $K$ en $\mathcal{P}K$ que asigna a cada $x \in K$, un conjunto $H_x \subseteq K$. Consideremos otra vez la aplicación $\varphi \in F$, definida de modo que, para cada $x \in K$, $\varphi(x) \neq f_x(x)$. $\varphi$ es la función característica de un determinado conjunto $G \subseteq K$. $G$ queda caracterizado por el hecho de que, para cada $x \in K$, $\varphi(x) = 1$, y, por ende, $f_x(x) = 0$. En otras palabras, $x \in G$ si y sólo si $x \notin H_x$. Así, pues, $G$ es el conjunto formado por los elementos de $K$ que no pertenecen al conjunto que respectivamente les asigna la inyección $gf$. Si $G = H_a$ para un cierto $a \in K$, tendremos que $a \in G$ si y sólo si $a \notin H_a = G$. Esta contradicción implica que $G$ cae fuera del alcance de la inyección $gf$, y que ésta no es una aplicación biyectiva. Como $gf$ es una inyección cualquiera, no hay una biyección de $K$ a $\mathcal{P}K$. Por lo tanto, $|K| < |\mathcal{P}K|$.

El Teorema de Cantor y su demostración fueron reformulados en términos similares a éstos por Bertrand Russell (1903, §§ 346–347). La principal diferencia reside en que Russell no habla de *conjuntos*, sino de *clases*, esto es, de extensiones de conceptos.[1] En la versión de Russell, el Teorema de

[1] "Una clase es . . . todos los objetos que satisfacen una función proposicional" (Whitehead y Russell, 1910–13, t. I, p. 23). El término 'función proposicional' se explica así: "Sea $\phi x$ un enunciado (*statement*) que contiene una variable $x$ y que se convierte en una proposición cuando se le asigna a $x$ cualquier significado determinado fijo. Entonces

Cantor dice, pues, que cualquier clase es menos numerosa que la clase de sus subclases.[2] En la misma obra, Russell nos cuenta que descubrió la paradoja que hoy lleva su nombre reflexionando justamente sobre las consecuencias del Teorema de Cantor. Dicha paradoja surge al considerar la clase de todas las clases que no son miembros de sí mismas. Si llamamos $C_R$ a esta clase russelliana, designamos con '$\{x: \varphi x\}$' la clase de todos los objetos $x$ que satisfacen la condición $\varphi x$, y simbolizamos con '$y \in z$' e '$y \notin z$' los enunciados '$y$ pertenece a la clase $z$' e '$y$ no pertenece a la clase $z$', tenemos que:

$$C_R = \{x: x \text{ es una clase y } x \notin x\}$$

Si $a$ es una clase cualquiera, es claro que $a \in C_R$ si y sólo si $a \notin a$. En particular, si la clase cualquiera considerada es la misma $C_R$, se obtiene la contradicción:

$$C_R \in C_R \leftrightarrow C_R \notin C_R$$

---

$\phi x$ se llama una 'función proposicional'" (p. 14). Cada objeto $a$ tal que la proposición $\phi a$ es verdadera "satisface" la función proposicional $\phi x$ y es, por ende, un miembro de la clase determinada por esta función proposicional. Aunque hay profundas diferencias entre *Principia Mathematica*, de donde tomo estas definiciones, y *The Principles of Mathematics* de 1903, en este punto concuerdan. Leemos en *The Principles* que "$\phi x$ es una función proposicional si, para todo valor de $x$, $\phi x$ es una proposición, determinada si $x$ es dado" (Russell 1903, § 22) y que "los valores de $x$ que hacen verdadera la función proposicional $\phi x$ [ . . . ] en general forman una *clase* y de hecho una clase puede definirse como todos los términos que satisfacen alguna función proposicional" (§ 23).

2 No es obvio para mí que el argumento que Russell toma de Cantor demuestre el teorema en su nueva versión. En efecto, si tomamos en serio la explicación lógico-lingüística de la noción de clase recogida en la nota 1, el número de subclases no vacías de una clase C no es mayor que el número de funciones proposicionales diferentes que satisfacen uno o más miembros de C. Si cada función proposicional es un enunciado, el número de las mismas dependerá de la índole del lenguaje que se usa. En particular, si el lenguaje es la "escritura conceptual" de Frege (1879), o la pasigrafía de Peano (1895 y ss.), o el sistema desarrollado más tarde por el mismo Russell (Whitehead y Russell, 1910–13), cualquiera de los cuales sólo admite oraciones de longitud finita con a lo sumo $\aleph_0$ predicados diferentes, el número total de enunciados diferentes que es posible hacer en él no es mayor que $\aleph_0$. A menos, claro está, que haya enunciados —específicamente, funciones proposicionales— inefables.

Esta contradicción es fatal para el sistema de fundamentación de la aritmética y el análisis elaborado por Frege en *Grundgesetze der Arithmetik* (1893/1903) y así lo reconoce el autor en un epílogo al tomo II (1903, pp. 252ss.), en el cual informa sobre la carta que Russell le envió el 16 de junio de 1902, cuando ese tomo ya estaba en prensa.[3] Me refiero a ese sistema en el Capitulo 2.2, la Sección 2.3.3 y los Apéndices X y XII. La contradicción subsiste, claro está, si '$y \in z$' significa, como de costumbre, '$y$ es un elemento del conjunto $z$' y pretendemos designar con $C_R = \{x\colon x \notin x\}$ el conjunto de todos los conjuntos que no son elementos de sí mismos. Pero puesta en esta forma no afecta la coherencia del pensamiento de Cantor, ya que éste no presupone que cada conjunto corresponda a un concepto, ni que cada concepto determine un conjunto.

Cantor deja muy en claro su posición a este respecto en una carta a Dedekind del 3 de agosto de 1899 (publicada en 1932; Cantor, GA, pp. 443ss.).[4] Presenta allí lo que hoy se conoce como la Paradoja de Burali-Forti, mas no como paradoja, sino simplemente como una indicación de que hay pluralidades, bien definidas mediante una caracterización verbal, que no forman una unidad, es decir, un conjunto. Al comienzo de la carta Cantor distingue dos especies de "pluralidad determinada (*bestimmte Vielheit*)".

> En efecto, una pluralidad puede ser de tal índole que el supuesto de que *todos* sus elementos "existen conjuntamente" (*die Annahme eines "Zusammenseins"* **aller** *ihrer Elemente*) lleva a una contradicción, de modo que es imposible captar esa pluralidad como una unidad, como "una cosa acabada". A tales pluralidades las llamo *pluralidades absolutamente infinitas* o *inconsistentes*. [ . . . ]

3 En su respuesta a Russell, el 22 de junio de 1902, Frege le dice que "su descubrimiento de la contradicción me ha sorprendido muchísimo y casi querría decir que me ha consternado, puesto que sacude el fundamento sobre el cual yo pensaba que podía construirse la aritmética" (Frege, WB, p. 213). En cambio, el joven Zermelo, que había descubierto la Paradoja de Russell un año antes, no la publicó y continuó sus investigaciones sobre la teoría de conjuntos sin hacerle caso, seguramente porque entendía que no era un obstáculo serio para dicha teoría. Véase Rang y Thomas 1980.

4 Zermelo la publicó en el apéndice de Cantor, GA, confundida con otra del 28 de julio, de la que provienen los cuatro primeros párrafos de la p. 443. La carta del 3 de agosto empieza en el quinto, con la frase "Gehen wir von dem Begriff einer bestimmten Vielheit . . . aus". Cf. Dauben 1979, p. 352; Grattan-Guinness 1974.

> En cambio, si la totalidad de los elementos de una pluralidad se deja concebir sin contradicción como "estando reunida" (*als "zusammenseiend"*), de modo que es posible captarla conjuntamente como "*una* cosa", la llamo *pluralidad consistente* o "conjunto" ("*Menge*").
>
> (GA, p. 443)[5]

Siguen tres asertos, en el estilo que se usará más tarde para enunciar axiomas de la Teoría de los Conjuntos:

> Dos pluralidades equinumerosas son o bien ambas "conjuntos", o bien ambas inconsistentes.
>
> Una pluralidad que es parte de un conjunto es un conjunto.
>
> Todo conjunto de conjuntos, cuando éstos se disuelven en sus elementos, es también un conjunto.
>
> (GA, p. 444)

El tercer aserto significa lo mismo que el Axioma V de Zermelo: "A cada conjunto $T$ corresponde un conjunto $\mathscr{S}T$ cuyos elementos son precisamente todos los elementos de los elementos de $T$" (Zermelo 1908a). La segunda alternativa mencionada en el primer aserto desconcierta a primera vista, ya que presupone que tiene sentido hablar de equinumerosidad entre pluralidades inconsistentes, aunque ninguna de ellas pueda considerarse como "una cosa acabada". Pero a la luz del uso que le da luego, me parece que todo lo que Cantor ha querido decir aquí es que si hay una pluralidad inconsistente y cada elemento discernible en ella puede ponerse en correspondencia exclu-

5 Las expresiones 'pluralidad consistente' y 'pluralidad inconsistente' (*konsistente*, bzw. *inkonsistente Mannigfaltigkeit*) habían sido utilizadas por Schröder en un libro publicado en 1890 (VAL, vol. I, p. 213) para referirse a pluralidades que podemos o, respectivamente, no podemos concebir como un todo (*als ein Ganzes denken*—p. 212). Como ejemplo de pluralidad inconsistente cita un par de proposiciones contradictorias, porque según él "el espíritu humano no es capaz de reunirlas". Sin embargo, cualquier demostración por reducción al absurdo tiene que *reunir* las conclusiones contradictorias que deriva de la hipótesis que intenta refutar, para concluir que dicha hipótesis es falsa. Es por cierto imposible concebir una situación que realice ("sea un modelo de") un par de proposiciones contradictorias, pero eso no impide concebir el todo formado por las proposiciones mismas.

siva con un elemento de otra pluralidad, y viceversa, entonces esta última también es inconsistente.

A continuación Cantor define los conceptos de 'orden simple' (nuestro 'orden lineal') y 'tipo de orden', y dice que una pluralidad simplemente ordenada se dice *bien ordenada* cuando cada una de sus partes tiene un primer elemento. Agrega que, en aras de la brevedad, llamará 'secuencia' ('Folge') a toda pluralidad bien ordenada, y observa que cada parte de una secuencia es también una secuencia. Define luego 'número ordinal' como el tipo de orden de un *conjunto* bien ordenado y pasa a considerar "el sistema de todos los ordinales", que designa con la letra Ω.[6] Recuerda haber demostrado que si $\alpha$ y $\beta$ son ordinales diferentes, o bien $\alpha < \beta$, o bien $\beta < \alpha$; que si $\alpha$, $\beta$ y $\gamma$ son ordinales tales que $\alpha < \beta$ y $\beta < \gamma$, entonces $\alpha < \gamma$, y que si Z es una colección de ordinales hay en Z un ordinal que precede a todos los otros. Concluye que el sistema Ω forma una secuencia ordenada por <. Por lo tanto, si Ω fuese un conjunto, su tipo de orden sería un determinado ordinal $\zeta$. Entonces $\zeta$, el tipo de orden de Ω, sería un ordinal posterior a todos los elementos de Ω. Pero $\zeta \in \Omega$, puesto que Ω es el sistema de todos los ordinales. Por lo tanto, $\zeta < \zeta$.[7] Esta contradicción se llama Paradoja de Burali-Forti, porque el matemático italiano la adujo —con otro propósito— en un artículo publicado en 1897.[8] Cantor infiere de ella, sencillamente, que el sistema de los ordinales no es un conjunto sino una pluralidad absolutamente infinita o inconsistente.

Enseguida nos invita a considerar el sistema de todos los alephs, que designa con la última letra del alfabeto hebreo, ת (tav). Como hay un aleph $\aleph_\alpha$ para cada $\alpha \in \Omega$, la inconsistencia de Ω implica que ת es inconsistente

6 Conviene quizás anotar que la expresión que usa Cantor para abreviar 'Ordnungszahl' ('número ordinal') es 'Zahl' ('número').

7 El razonamiento resulta quizás más claro si invocamos la Proposición 4 del Apéndice II: Un conjunto bien ordenado no puede ser isomorfo a uno de sus segmentos. Pero si el ordinal $\zeta \in \Omega$ es el tipo de orden de Ω, Ω es isomorfo a $[\zeta]$, el segmento de Ω determinado por $\zeta$. El lector a quien, con toda razón, le moleste la idea de que una pluralidad inconsistente se diga bien ordenada puede reordenar así el razonamiento de Cantor: Suponga primero que Ω es un conjunto; muestre que si lo es, está bien ordenado por <. Por lo tanto, Ω tiene un tipo de orden $\zeta \in \Omega$ y es isomorfo al segmento $[\zeta] \subseteq \Omega$. Como esto es imposible si Ω está bien ordenado, Ω no es un conjunto.

8 Burali-Forti 1897. Explico el argumento de Burali-Forti en el Apéndice IV.

(en virtud del primero de los tres asertos iniciales de Cantor arriba citados). Cantor basa en esta conclusión un argumento que según él demuestra que el cardinal de un conjunto infinito siempre es un aleph: Sea $V$ una pluralidad infinita cuya numerosidad no es un aleph. Es claro, entonces, que, cualquiera que sea el ordinal $\alpha$, si $f_\alpha$ es una inyección de $[\alpha]$ en $V$, $f_\alpha$ no es una biyección. De esto se sigue, según Cantor, que $\Omega$ es equinumeroso a una pluralidad $V' \subseteq V$.[9] Por lo tanto, en virtud de los dos primeros asertos iniciales de Cantor citados arriba, $V'$ y $V$ son pluralidades inconsistentes y $V$ no puede ser un conjunto. Si el argumento precedente es válido, constituye una demostración fácil del Teorema del Buen Orden: toda pluralidad cuyo cardinal sea un aleph se deja biyectar en un segmento de $\Omega$ y bien ordenar por ésta, y una pluralidad que no se deja bien ordenar, lisa y llanamente no es un conjunto.[10] Esta demostración —"por exclusión de lo disconforme"— encierra un peligro. Como nadie sabría coordinar los puntos de un trazo con un segmento de $\Omega$, se puede pensar —o decidir— que ellos no constituyen un conjunto. Este diagnóstico afectaría asímismo al sistema $\mathbb{R}$ de los números reales y a todas las pluralidades equinumerosas con él. Para eliminar

9 Cantor no explica cómo llega a esta conclusión. Pero he aquí un modo como pudiera haberla defendido. Digamos que una aplicación $g$ está incluida en una aplicación $h$ (simbólicamente, $g \subseteq h$) si el dominio de $g$ está incluido en el dominio de $h$ y $g(x) = h(x)$ para cada $x$ en el dominio de $g$. Sea $V$ una pluralidad infinita cuyo cardinal no es un aleph. Como $V$ es infinito, es claro que hay una inyección $f_\omega$: $[\omega] \to V$. Como el cardinal de $V$ no es $\aleph_0$, es claro que $f_\omega$ no es biyectiva y que hay por lo menos un elemento de $V$ que no está en el alcance de $f_\omega$. Partiendo de $f_\omega$ puede establecerse una secuencia transfinita de aplicaciones $f_\alpha$: $[\alpha] \to V$ tales que, para cualesquiera ordinales transfinitos $\alpha$ y $\beta < \alpha$, (i) $f_\beta \subseteq f_\alpha$, (ii) $f_\alpha$ es inyectiva y (iii) $f_\alpha$ no es biyectiva. Sea $\alpha > \omega$ y supongamos que $f_\beta$ está definida y satisface las condiciones (i)–(iii) para todo ordinal transfinito $\beta < \alpha$. Si $\alpha$ es un ordinal sucesor, hay un $\beta \geq \omega$ tal que $\alpha = \beta + 1$, y hay un elemento $v_\beta \in V$ que no está en el alcance de $f_\beta$. $f_\alpha$: $[\beta + 1] \to V$ se define así: si $\gamma < \beta$, $f_\alpha(\gamma) = f_\beta(\gamma)$; $f_\alpha(\beta) = v_\beta$. Si $\alpha$ es un ordinal límite, $f_\alpha$ se define así: para cada ordinal transfinito $\beta < \alpha$, $f_\alpha|[\beta] = f_\beta$ ($f_\alpha$ coincide con $f_\beta$ en el dominio de esta última). Estas definiciones aseguran que $f_\alpha$ cumple en ambos casos las condiciones (i) y (ii). Es obvio que también cumple la condición (iii), ya que de otro modo el cardinal de $V$ sería un aleph. La correspondencia $\alpha \mapsto f_\alpha(\alpha)$ entre $\Omega$ y la pluralidad $\{f_\alpha(\alpha):\alpha \in \Omega\} = V' \subset V$ es biunívoca.

10 Se sabe que en 1896 ó 1897 Cantor envió a Hilbert una demostración del Teorema del Buen Orden. No conocemos su tenor, pero G. H. Moore (1982, p. 51) conjetura que se basaba en el mismo argumento utilizado en la carta a Dedekind del 2 de agosto de 1899. Hilbert no se dejó convencer.

esa opción —ruinosa para el proyecto de fundar el análisis en una teoría de conjuntos— Cantor tendría que haber suplementado su demostración del Teorema del Buen Orden con una prueba de la Hipótesis del Continuo.

Zermelo, que publicó dos pruebas bastante más elaboradas del Teorema del Buen Orden en 1904 y 1908, levantó en 1932 dos objeciones contra el argumento de Cantor: En primer lugar, la biyección de Ω en una parte de V que Cantor invoca pero no define, tendría, al parecer, que definirse mediante una secuencia transfinita de selecciones sucesivas, lo cual significaría aplicar nuestra intuición del tiempo a un proceso que rebasa toda intuición. En segundo lugar, "la demostración opera con pluralidades 'inconsistentes', y hasta posiblemente con conceptos contradictorios, y ya por eso sería lógicamente inadmisible" (Cantor, GA, p. 451 n. 1). La primera objeción es sin duda correcta.[11] En sus demostraciones del Teorema del Buen Orden, Zermelo la elude gracias a su famoso Axioma de Selección, que autoriza la selección *simultánea* de un número infinito de objetos (*vide* Capitulo 1.7). La segunda objeción, en cambio, me parece confusa. Contradictorio y, por ende, lógicamente inadmisible, sería suponer que alguna de las pluralidades inconsistentes mencionadas en la prueba de Cantor es una unidad, esto es, un conjunto, pero no veo que su razonamiento envuelva tal suposición.

A Zermelo le ha inquietado quizás la liviandad con que Cantor designa mediante una sola letra (Ω, *V*) a pluralidades que juzga imposible concebir como unidades. Por la misma índole de tales pluralidades no está claro de qué manera o en qué preciso sentido puede uno referirse a ellas con un nombre o un pronombre. A primera vista parecería que una tal referencia no es posible: si la existencia conjunta de los elementos de una cierta pluralidad implica contradicción una denominación que pretendidamente los abarque a todos simplemente no designa nada. Sin embargo, es claro que una frase castellana como "los ordinales" se refiere a *cada* ordinal, aunque sea imposible referirse a *todos*. Por ejemplo, si asevero que entre los ordinales no hay un máximo, digo, sin ambigüedad ni oscuridad alguna, que para *cada* ordinal hay *otro* que le sigue. En general, si el predicado 'es un *P*' distingue ciertos objetos, puede usárselo en expresiones nominales tales como 'los *P*' o '{*x*:*x* es un *P*}' para denotar a esos objetos en su disgregada pluralidad

[11] Cf. la construcción explicada en la nota 9. Obsérvese que envuelve la selección sucesiva de objetos $v_\alpha \in$ V, uno para cada $\alpha \in \Omega$.

aunque no se dejen congregar en un conjunto. Diré que la referencia a esos objetos en tal caso es *distributiva*, pero no *colectiva*. Llamaré 'designador distributivo' a una expresión utilizada para referirse distributivamente a ciertos objetos.[12] Claro está que la referencia distributiva inevitablemente confiere una suerte de unidad a la pluralidad referida. Se corre, entonces, el riesgo de tratar a ésta *de palabra* como lo que no es *de veras*: un objeto susceptible de agruparse con otros en nuevas unidades. En la Sección 1.8.4 consideraremos el procedimiento adoptado por von Neumann en la década de 1920 para obviar este riesgo, y que consiste en tratar a las pluralidades que Cantor llamó inconsistentes como un tipo de objetos, sí, pero inconfundible con los objetos ordinarios y en particular con los conjuntos, por cuanto en una aseveración de la forma '$x \in y$' ('$x$ es un elemento de $y$') un objeto de ese tipo sólo puede ser denotado por $y$, nunca por $x$.

El reconocimiento de que no toda pluralidad bien definida, esto es, caracterizada mediante un criterio diagnóstico preciso, es un conjunto en el sentido de Cantor no es fatal para su programa, pero suscita un problema que marcará todo el ulterior desarrollo del mismo por otros autores. Cantor alude a él ya en una tercera carta a Dedekind, del 28 de agosto de 1899. Específicamente, se pregunta cómo puede uno saber que las pluralidades bien ordenadas a las que él asigna los cardinales $\aleph_0, \aleph_1, \ldots, \aleph_\omega, \ldots \; \aleph_{\omega_1}, \ldots$ son efectivamente conjuntos. "¿No pudiera pensarse que ya *estas* pluralidades son 'inconsistentes', sólo que la contradicción inherente en suponer la 'existencia conjunta de todos sus elementos' *todavía no se ha hecho notar*?" (GA, p. 447). Su respuesta es tan audaz como profunda:

> La pregunta debe extenderse también a las pluralidades finitas y un examen detenido conduce al resultado siguiente: aun en el caso de las pluralidades finitas *no* cabe dar una "prueba" de su "consistencia". En otras palabras: el hecho de la "consistencia" de las pluralidades finitas es una verdad simple, indemostrable, es "*el axioma* de la aritmética" (en el antiguo sentido del término). Y del mismo modo, la "consistencia" de las pluralidades a las que asigno los alephs como cardinales es "el axioma de la ampliada aritmética transfinita".
>
> (Cantor, GA, pp. 447–48)

[12] Por ejemplo, la frase 'universo del discurso', empleada en semántica para referirse a todo cuanto pueda ser tema de conversación en el lenguaje bajo estudio, es ciertamente un designador distributivo si el lenguaje en cuestión es natural.

Esta respuesta tiene dos aspectos que consideraré sucesivamente: (A) La duda señalada afecta a la aritmética finita no menos que a la transfinita. (B) El matemático tiene la libertad de superar tales dudas postulando axiomas.

(A) En un escrito anterior Cantor ya había observado que las limitaciones humanas que alegadamente impiden o entraban la matemática del infinito dificultan asímismo la consideración de cantidades finitas —al menos cuando son muy grandes— y que los recursos de que la matemática se ha valido tradicionalmente para superar tales limitaciones le sirven también para pensar sobre lo infinito. El 28 de febrero de 1886 Cantor dirigió una carta al médico berlinés, Dr. A. Eulenburg, sobre el "horror del infinito" común entre médicos y hombres de ciencia. Luego la publicó, con otras sobre temas afines, en una revista filosófica, bajo el título de "Comunicaciones sobre la doctrina del transfinito" (Cantor 1887/88). La versión impresa incluye una larga nota sobre el infinito matemático en la obra de San Agustín, Orígenes y Santo Tomás de Aquino. Cantor cita *in extenso* el Capítulo 19 del Libro XII de *La Ciudad de Dios*, donde San Agustín refuta a quienes sostienen que las cosas "que son infinitas no pueden ser comprendidas ni siquiera por el conocimiento divino". No puede ser, dice San Agustín, que Dios no conozca todos los números. Ahora bien, "cada número está determinado por sus propiedades, de modo que ninguno de ellos puede ser igual a otro. Por lo tanto, son desiguales y diversos entre sí, y aunque cada uno es finito, todos son infinitos (*et singuli quique finiti sunt, et omnes infiniti sunt*)." Puesto que lo que el conocimiento comprende tiene que "estar definido (*finitur*) por la comprensión de quien conoce", la infinidad de los números "está definida (*finita est*) para Dios de cierta manera inefable (*quoddam ineffabile modo*)". Según Cantor, cuando San Agustín afirma que Dios capta totalmente, en forma intuitiva, la multitud infinita de los números, "reconoce al mismo tiempo este conjunto formalmente como un todo infinito actual, como un *Transfinitum*, y nos vemos compelidos a seguirle en esto" (GA, p. 402). Se objetará quizás que, aunque estemos forzados a admitir que el conjunto $\mathbb{N}$ de los enteros positivos es un objeto infinito, no nos es lícito entrar a considerar su ordinal $\omega$ ni su cardinal $\aleph_0$, por cuanto "*nosotros*, en nuestra esencial limitación, no somos capaces de captar con un solo golpe de vista (*uno intuitu*) todos los infinitos números individuales $n$ que pertenecen al conjunto $\mathbb{N}$." Contra esta objeción, Cantor aduce que nadie es capaz de "representarse *uno intuitu*

distinta y precisamente” todas las unidades comprendidas en un número finito bastante grande, por ejemplo, mil millones.

> Y sin embargo tenemos derecho a considerar los números finitos, por grandes que sean, como objetos del conocimiento discursivo humano, e investigarlos científicamente según su modo de ser. El mismo derecho nos compete con respecto a los números transfinitos.
>
> [ . . . ]
>
> Rara vez poseemos un concepto del que se pueda decir que es un “conceptus rei proprius ex propriis”, con el que captamos y conocemos adecuadamente una cosa tal como es en y por sí misma, sin ayuda de la negación, de un símbolo o de un ejemplo. Antes bien, en nuestro conocer dependemos casi siempre de un “conceptus proprius ex communibus” que nos capacita para determinar una cosa mediante predicados generales y con ayuda de comparaciones, exclusiones, símbolos o ejemplos, de suerte que se distinga bien de cualquier otra cosa. Compárese, por ejemplo, el método [explicado en el Capítulo 1.3] que usé para definir las magnitudes numéricas *irracionales*. Iré *al extremo de aseverar incondicionalmente* que este segundo modo de determinación y delimitación de las cosas es *incomparablemente más sencillo, más cómodo* y *más fácil* en el caso de los *números transfinitos pequeños* (por ejemplo, $\omega$, u $\omega + 1$, u $\omega^{\nu}$, si $\nu$ es un entero finito pequeño) que tratándose de números finitos muy grandes, en cuyo caso, sin embargo, también hemos de recurrir al mismo medio auxiliar ajustado a nuestra naturaleza imperfecta.
>
> (Cantor, GA, pp. 402–3)

(B) La libertad del matemático es el tema de lo que, a mi modo de ver, es el texto filosófico más importante que nos ha dejado Cantor. Constituye el § 8 del escrito Nº 5 de la serie “Sobre variedades lineales infinitas de puntos”, que ya he citado varias veces. Empieza distinguiendo dos sentidos en que cabe hablar de la existencia (*Existenz*) o realidad actual (*Wirklichkeit*) de los números enteros, finitos o infinitos. Por una parte, podemos considerar que ellos existen

> en cuanto ocupan, sobre la base de definiciones, un lugar enteramente determinado en nuestro entendimiento, se distinguen perfectamente de todos los demás componentes de nuestro pensamiento, tienen determinadas relaciones con ellos, y así modifican de una determinada manera la sustancia de nuestro espíritu. Permitáseme llamar esta forma de realidad (*Realität*)

> de nuestros números su *realidad intrasubjetiva* o *inmanente*.
> (Cantor, GA, p. 181)

Por otra parte, en cuanto las distintas clases de números, (I), (II), (III), etc., son representantes de numerosidades que efectivamente se encuentran en la naturaleza,

> cabe también atribuir a los números realidad actual (*Wirklichkeit*) porque hay que verlos como una expresión (*Ausdruck*) o reflejo (*Abbild*) de sucesos y relaciones en el mundo exterior que enfrenta al intelecto. [ . . . ] Llamo esta segunda forma de realidad la *realidad transubjetiva* o *trascendente*.
> (Cantor, GA, p. 181)

Ahora bien, la matemática, "en la elaboración de su caudal de ideas, tiene que considerar *única* y *exclusivamente* la realidad *inmanente* de sus conceptos, y no tiene *ninguna* obligación de examinarlos en lo que respecta a su realidad *trascendente*" (GA, p. 182). Cantor no piensa, sin embargo, que cualquier concepción arbitraria, con tal que no genere contradicciones, tiene derecho a un lugar en el sistema de las ideas matemáticas. "La matemática es completamente libre en su desarrollo", pero éste está sujeto a *dos* restricciones: los conceptos matemáticos deben estar libres de contradicción interna, y *además* deben "estar en relaciones firmes, ordenadas mediante definiciones, con los conceptos previamente formados, ya existentes y probados" (GA, p. 182).[13] Esta doble exigencia "deja al arbitrio un campo muy reducido". Más aún, "cada concepto matemático lleva en sí el correctivo necesario: si no es idóneo y fecundo, él mismo lo demuestra muy pronto con su inutilidad, y es entonces descartado por falta de éxito." En cambio, según Cantor,

> cualquier estrechamiento superfluo del afán de investigación matemática envuelve un peligro mucho mayor, tanto más grande cuanto que no puede

13 Cantor prosigue: "En particular, al introducir nuevos números [la matemática] sólo está obligada a dar de ellos definiciones que les confieran una tal determinación y, eventualmente, una tal relación con los números más antiguos, que, dado el caso, se puedan distinguir entre ellos de un modo determinado. En cuanto un número satisface todas estas condiciones se puede y se debe considerarlo como existente y real (*existent und real*) en la matemática" (GA, p. 182).

> dársele ninguna justificación basada en la esencia de la ciencia; pues la *esencia* de la *matemática* consiste justamente en su *libertad*.
>
> (Cantor, GA, p. 182)

Con todo, en el mismo § 8 en que defiende tan rotunda e inequívocamente los derechos de una matemática autónoma, basada en su propia inventiva y fiel sólo a su propia historia, Cantor declara su convicción personal de que la realidad inmanente y la realidad trascendente siempre concuerdan, en cuanto

> un concepto que deba caracterizarse como existente en el primer sentido siempre posee también una realidad trascendente en ciertos —y en verdad, en infinitos— respectos, cuya comprobación, claro está, es en general una de las tareas más laboriosas y difíciles de la metafísica y suele tener que diferirse hasta que el desarrollo natural de una de las otras ciencias descubra el significado transcendente del concepto en cuestión.
>
> (Cantor, GA, p. 181)[14]

Esta creencia de Cantor hace comprensible —aunque en mi parecer no justifica— dos epígrafes latinos que puso a la cabeza de su obra más madura, "Aportes a la fundamentación de la teoría de los conjuntos transfinitos" (1895/97), a saber, la frase de Newton, "hypotheses non fingo" ("no forjo hipótesis"), y otra, tomada de Francis Bacon, según la cual "no le damos leyes al intelecto y a las cosas según nuestro arbitrio, sino que como escribas fieles anotamos y transcribimos las que anuncia y pronuncia la voz de la propia naturaleza."[15] Un tercer epígrafe, tomado de la Biblia, evoca la observación, arriba citada, de que la comprobación de la realidad trascendente de los conceptos matemáticos debe dejarse al trabajo futuro de otras

14 Cantor agrega en una nota que esta convicción suya concuerda con los principios de la filosofía de Platón y "con un rasgo central del sistema de Spinoza", el expresado en la Prop. VII de la Parte II de la *Ética*: "El orden y la conexión de las ideas es el mismo que el orden y la conexión de las cosas"

15 Dauben (1979), pp. 236–39, explica que con las expresiones citadas Cantor buscaba sobre todo manifestar su rechazo de los infinitésimos y la geometría no-arquimédica de Veronese (¿así entendía la libertad de la matemática?). Dauben señala que en una carta a Vivanti del 13 de diciembre de 1893, en la que ya aducía —contra Veronese— la cita de Bacon, Cantor equipara "la naturaleza" con "lo posible". A la luz de la referencia a Spinoza en Cantor 1883 (v. nota 14), me parece que esta equiparación no significa, como sugiere Dauben, que lo que aquí se llama "la naturaleza" no es más que "lo

disciplinas: “Digo cosas que ahora están ocultas, pero llegará el tiempo en que una persistente diligencia las saque a la luz del día.”

La polémica apasionada de Cantor en la década de 1880 contra quienes, desde Aristóteles, han negado el infinito actual, puede entenderse como un primer paso hacia la certificación de la realidad trascendente de los conceptos inventados por él, pero también como un intento para demostrar informalmente su consistencia.[16] En una carta a G. Eneström del 4 de noviembre de 1885 (Cantor 1886), distingue tres modos como puede considerarse —y cuestionarse— el infinito actual: (i) en cuanto es “lo absoluto” en Dios eterno omnipotente y extramundano o *natura naturans*; (ii) en cuanto es “lo transfinito”, presente en concreto en la *natura naturata*, y (iii) en cuanto puede ser captado en abstracto por el conocimiento humano “en la forma de números actualmente infinitos —o transfinitos, como los he llamado— o en la forma aún más general de los *tipos de orden transfinitos*” (GA, p. 372). Dejando de lado el primer modo, Cantor señala que los otros dos dan lugar a cuatro posiciones filosóficas diferentes: algunos, como Cauchy, Gauß, León XIII (en la encíclica *De philosophia christiana*) y “todos los llamados *positivistas*”, niegan el infinito actual en concreto y en abstracto; otros, como Descartes, Spinoza, Leibniz, Locke, Lotze, lo sostienen en concreto y lo niegan en abstracto; algunos neoescolásticos lo niegan en concreto pero lo afirman en abstracto; por último,

> el infinito actual puede ser *afirmado* tanto en concreto como en abstracto; pocos toman esta posición, que considero la *única correcta*; quizás sea yo el primero que representa este punto de vista con plena decisión y en todas sus consecuencias, pero estoy seguro de que no seré el último en defenderlo.
>
> (Cantor, GA, p. 373)

---

posible”, sino que “lo posible” es nada menos que “la naturaleza”: para Cantor, como para nuestro contemporáneo David Lewis, lo posible es lo realmente real.

16 Cantor 1883, 1886, 1887/88 (GA, pp. 165ss., 370ss., 379ss.). Dauben 1979, Capítulo 6, resume y comenta muy bien “la filosofía del infinito” contenida en estos escritos.

Según Cantor, el error de los positivistas consiste en que de antemano atribuyen “o más bien imponen” a todos los números las propiedades de los números finitos, “*mientras que los números infinitos, si en absoluto han de ser concebibles de algún modo, tienen que constituir, por su oposición a los números finitos, una clase de números totalmente nueva, cuya índole depende enteramente de la naturaleza de las cosas y es objeto de investigación, no de nuestro arbitrio o de nuestros prejuicios*” (GA, p. 372).[17] Así, todos los argumentos de Aristóteles contra el infinito actual (por ejemplo, en *Metaph.* K, 10) se basan en el supuesto de que no hay más que números finitos, supuesto que Aristóteles infería, a su vez, de que sólo había conocido enumeraciones de conjuntos finitos (GA, p. 174). De un modo análogo, Santo Tomás de Aquino niega que pueda haber una multitud actualmente infinita, por cuanto toda multitud es una multitud de una cierta especie, las especies de multitud corresponden a las especies de números, y ninguna de éstas es infinita, ya que todo número es una multitud medida por el uno (*Summa theol.* I, q. 7, art. 4); pero este argumento desconoce que “en lo transfinito está presente y en cierto modo almacenada una riqueza mucho mayor de formas y de ‘species numerorum’ que en el campo relativamente pequeño de lo finito ilimitado” (GA, p. 404). Por último, el enfático rechazo del infinito actual por grandes matemáticos del siglo XIX como Cauchy y Gauß se explica en parte, según Cantor, como una reacción contra el uso y abuso de lo infinitamente pequeño en la matemática del siglo XVIII —que también Cantor repudia— pero también debe atribuirse a “una especie de *miopía*, que nos roba la posibilidad de ver el infinito actual, a pesar de que en [la forma de] su portador supremo y absoluto nos creó y nos conserva y en sus formas secundarias, transfinitas, nos rodea por todas partes y mora en nuestro espíritu” (GA, pp. 374–75).

17 Trascribo este pasaje, destacado en el original por el autor, porque Michael Hallett, en un libro por lo demás bastante instructivo, hace gran caudal de lo que llama “el principio del finitismo de Cantor”, que formula así: “Lo transfinito está a la par con lo finito y matemáticamente se lo debe tratar, en lo posible, igual que a lo finito (*like the finite*)” (Hallett 1984, p. 7).

## 1.7 EL TEOREMA DEL BUEN ORDEN Y EL AXIOMA DE SELECCIÓN

"Aportes a la fundamentación de la teoría de los conjuntos transfinitos" (1895/97) es la última publicación importante de Cantor. Deja sin resolver los dos problemas de cuya solución —como vimos en el Capitulo 1.5— depende la alegada suficiencia de la aritmética transfinita cantoriana para medir todas las multitudes del universo: el problema del continuo y el problema del buen orden. Cantor no tuvo éxito en sus intentos por demostrar la Hipótesis del Continuo y no logró persuadir a Hilbert con su demostración del Teorema del Buen Orden (Capitulo 1.6, nota 10). En una famosa conferencia ante el Segundo Congreso Internacional de Matemáticos (París, 1900), Hilbert planteó 23 problemas que la matemática del siglo XX debía encarar. El primero de todos es "el problema de Cantor de la cardinalidad del continuo". Las investigaciones de Cantor sobre "los conjuntos de números reales ordinarios o puntos" sugieren un teorema que Hilbert considera muy plausible: Todo conjunto infinito de números reales es equinumeroso ya sea con el conjunto de los enteros positivos $\{1, 2, \ldots\}$, ya sea con el conjunto de todos los reales, es decir, con el continuo mismo. A este propósito, Hilbert menciona otra tesis de Cantor que quizás encierre la clave de una demostración del teorema anterior: El conjunto de todos los números reales se deja ordenar bien.

> Me parece sumamente deseable obtener una prueba directa de este notable aserto de Cantor, tal vez indicando efectivamente un ordenamiento de los números tal que dentro de cada subconjunto pueda señalarse un primer número.
>
> (Hilbert, GA, III, p. 299)

El Tercer Congreso Internacional de Matemáticos se reunió en Heidelberg en el verano de 1904. El 10 de agosto, Julius König presentó un argumento que demostraba, según él, que el cardinal del continuo no es un aleph y que, por consiguiente, el conjunto de los números reales no puede ser bien ordenado. Esta conclusión se deduce de la igualdad

$$\aleph_\alpha^{\aleph_0} = \aleph_\alpha \cdot 2^{\aleph_0} \quad (1)$$

(donde $\alpha$ es un ordinal cualquiera), que König toma de la disertación doctoral completada en 1901 bajo la tuición de Hilbert por el discípulo de Cantor, Felix Bernstein. Es fama que Cantor quedó consternado después de escuchar la ponencia de König; pero esa misma noche Zermelo detectó una falla en su razonamiento y la comunicó al Congreso al otro día: la demostración de la igualdad (1) ofrecida en la disertación de Bernstein no vale con toda la generalidad requerida por König, sino únicamente si $\alpha < \omega$ (y de hecho Bernstein mismo usa (1) sólo en este caso).[1] Seis semanas después, en una carta fechada el 24 de septiembre, Zermelo hizo llegar a Hilbert la prueba del Teorema del Buen Orden que bosquejaré a continuación.

Como Zermelo (1904) señala al final de su prueba, ella depende esencialmente del siguiente supuesto:

**AS** Si $\mathcal{K}$ es una colección de conjuntos no vacíos y $K$ es el conjunto de todos los elementos pertenecientes a los conjuntos de la colección $\mathcal{K}$ ($K = \bigcup\mathcal{K}$), hay una aplicación $f{:}\mathcal{K} \to K$ tal que, para cada $k \in \mathcal{K}$, $f(k) \in k$.

Como $f$, por así decir, "selecciona" un elemento de cada conjunto de la colección $\mathcal{K}$, diré que $f$ es un selector para $\mathcal{K}$ y llamaré a $f(k)$ el elemento selecto del conjunto $k$. Para probar que un conjunto cualquiera $M \neq \varnothing$ se deja ordenar bien, Zermelo considera un selector para $\mathcal{P}M\setminus\{\varnothing\}$, al que designa con la letra $\gamma$ ($\gamma$ selecciona un elemento en cada parte no vacía de $M$). Diremos, con Zermelo, que una parte $G$ de $M$ es un $\gamma$-conjunto si $G$ está bien ordenada por una relación $<_G$ y, para $g \in G$, $g = \gamma(M\setminus\{x \in G: x <_G g\})$ ($g$ es el elemento seleccionado por $\gamma$ en el complemento en $M$ del

[1] König 1905 es una versión corregida de esta ponencia. Allí König admite que sólo ha demostrado la proposición condicional: "Si el lema (1) de Bernstein vale para todo ordinal $\alpha$, entonces la Hipótesis del Continuo es falsa". Como, por otra parte, la falsedad de la hipótesis del continuo implica el lema (1), la negación de este lema puede tomarse como una formulación alternativa de la Hipótesis del Continuo que los partidarios de ésta podrían ensayarse en demostrar. König 1905a combate la posibilidad de bien ordenar el continuo con otro argumento.

conjunto de los predecesores de $g$ en $\langle G,<_G\rangle$). Diremos asímismo que un elemento $g \in M$ es un γ–elemento si $g$ pertenece a un γ–conjunto. Así, por ejemplo, es claro que el elemento selecto de $M$, $\gamma(M)$, es un γ–elemento, puesto que $\{\gamma(M)\}$ reúne los requisitos para ser un γ–conjunto (está bien ordenado, en un sentido trivial, y su único elemento es precisamente el elemento seleccionado por γ en el complemento del conjunto ∅ de sus predecesores). Como un γ–conjunto es un conjunto bien ordenado, su tipo de orden es un ordinal. Sean α y β los tipos de orden de los γ–conjuntos $\langle G_\alpha,<_\alpha\rangle$ y $\langle G_\beta,<_\beta\rangle$, respectivamente. Si $\alpha < \beta$, hay un morfismo inyectivo $\varphi$: $G_\alpha \to G_\beta$ (Apéndice II, Proposición 5). Supongamos que $\varphi(x) \neq x$ para algún $x \in G_\alpha$. Entonces $\{x \in G_\alpha : x \neq \varphi(x)\}$ es una parte no vacía del conjunto bien ordenado $G_\alpha$ y tiene, por ende, un primer elemento $g_1$. Entonces $U = \{x \in G_\alpha : x <_\alpha \varphi(x)\} = \{\varphi(x) : x \in G_\alpha \wedge \varphi(x) <_\beta \varphi(g_1)\} = V$ y, por definición, $g_1 = \gamma(M \setminus U) = \gamma(M \setminus V) = \varphi(g_1)$. Vemos así que $\{x \in G_\alpha: x \neq \varphi(x)\} = \varnothing$ y que $\varphi(x) = x$ para cada $x \in G_\alpha$. De suerte que $G_\alpha \subseteq G_\beta$ y, como $f$ es un morfismo, la relación $<_\alpha$ concuerda en $G_\alpha$ con $<_\beta$. Llamaré $M_\gamma$ al conjunto de todos los γ–elementos de $M$ (Zermelo lo llama $L_\gamma$, pero no veo la ventaja de introducir una letra nueva para este propósito). Sabíamos ya que $M_\gamma \neq \varnothing$. Ahora hemos aprendido que $M_\gamma$ es la unión de una secuencia, posiblemente transfinita, de γ–conjuntos $G_1, G_2, \ldots$ tales que $G_1 \subseteq G_2 \subseteq \ldots$ Obsérvese, en particular, que $G_1 = \{\gamma(M)\}$, de modo que $\gamma(M)$ figura —como primer elemento— en cada γ–conjunto. Sean $a$ y $b$ dos γ–elementos y sea α el primer ordinal tal que $\{a,b\} \subseteq G_\alpha$. Digamos que $a < b$ si $a <_\alpha b$. Es claro que $<$ determina un orden lineal en $M_\gamma$. Este orden es un buen orden. Para comprobarlo, consideremos un conjunto no vacío $A \subseteq M_\gamma$ y un $a \in A$. Sea α el primer ordinal tal que $a \in G_\alpha$. Entonces $\{x \in M_\gamma : x \leq a\} \subseteq G_\alpha$ y si $B = A \cap \{x \in M_\gamma : x \leq a\}$, $<$ concuerda con la relación de buen orden $<_\alpha$ en $B \subseteq G_\alpha$. Por lo tanto, $\langle B,<\rangle$ tiene un primer elemento, que es también el primer elemento de $\langle A,<\rangle$. Vemos así que cualquier parte no vacía de $M_\gamma$ tiene un primer elemento en el orden $<$. Por lo tanto $\langle M_\gamma,<\rangle$ es un conjunto bien ordenado. Finalmente, probaré que $M_\gamma = M$. Sea $M^* = M \setminus M_\gamma$. Si $M^* \neq \varnothing$, podríamos extender el orden $<$ a $M_\gamma \cup \{\gamma(M^*)\}$ estipulando que para todo $x \in M_\gamma$, $x < \gamma(M^*)$. Entonces, evidentemente, $M_\gamma \cup \{\gamma(M^*)\}$ sería un γ–conjunto y $\gamma(M^*) \in M_\gamma$. Como esto es imposible, dada la definición de $M^*$, concluimos que $M^* = \varnothing$ y que $M$ es idéntico al conjunto bien ordenado $M_\gamma$. Como $M$ es un conjunto

cualquiera, todo conjunto admite un buen orden.

El supuesto **AS** que Zermelo destaca expresamente al final de su prueba se conoce desde entonces como el Axioma de Selección (alemán, *Auswahlsaxiom*; francés, *Axiome du Choix*; inglés, *Axiom of Choice*).[2] Dice Zermelo al respecto: "Este principio lógico no puede derivarse de otro más simple, pero se aplica universalmente sin titubeos en el razonamiento matemático" (1904, p. 516). Vemos así que para establecer como un teorema lo que Cantor había llamado "una ley del pensamiento" Zermelo tiene que invocar un "principio lógico" que no figura en los escritos de Aristóteles, ni en las obras más recientes de Boole o de Frege. No sabría decir si el Axioma de Selección es más simple que el Teorema del Buen Orden. Lo que sí es claro es que la adopción de cualquiera de los dos como principio que no se demuestra permite demostrar al otro. El Teorema del Buen Orden se infiere del Axioma de Selección, como hemos visto, con un argumento sutil pero no difícil. El Axioma de Selección se deduce del Principio del Buen Orden de un modo brutalmente fácil: Sea $\mathcal{K}$ una colección de conjuntos no vacíos y $<$ un buen orden en $K = \bigcup\mathcal{K}$. Entonces $k \in \mathcal{K}$ sólo si $k \subseteq K$ y evidentemente hay un selector $f\colon \mathcal{K} \to K$ que asigna a cada $k \in \mathcal{K}$ un valor $f(k) \in k$, a saber, la aplicación que asigna a $k$ su primer elemento en $\langle K,<\rangle$.

La demostración del Teorema del Buen Orden por Zermelo (1904) causó gran revuelo entre los especialistas y reacciones muy adversas, especialmente en Francia. Los ilustres matemáticos Borel y Lebesgue, que habían dado táctiamente por supuesto el Axioma de Selección en trabajos anteriores, lo declararon inadmisible cuando vieron que podía invocárselo para establecer la conclusión, a juicio suyo inverosímil, de que el continuo se deja ordenar bien. No puedo entrar aquí en los detalles de esta polémica, que G. H. Moore analiza lúcidamente en su magistral historia del Axioma de Selección (1982).

En respuesta a sus críticos, Zermelo produjo la "Nueva demostración de la posibilidad de un buen orden" (1908a). El artículo está dividido en dos partes que contienen la nueva demostración y un examen de las objeciones contra la demostración anterior, respectivamente. La segunda demostración depende del Axioma de Selección, como la primera, pero difiere de ella en cuanto ahora Zermelo evita cuidadosamente invocar resultados de la teoría

[2] Algunos autores de habla castellana prefieren decir 'Axioma de Elección'.

cantoriana de los ordinales. Antes bien, al mostrar cómo se puede introducir un buen orden en cualquier conjunto $M$ —dado un selector para $\mathcal{P}M\setminus\{\varnothing\}$ —Zermelo ilumina la estructura subyacente a esa teoría y prepara el terreno para su reformulación puramente conjuntista en términos de las relaciones $\in$ y $\subseteq$.

La demostración va precedida del enunciado explícito de dos supuestos no demostrados: (I) Todos los elementos de un conjunto $M$ que cumplen una condición bien definida $E$ forman un subconjunto (*Untermenge*) $M_E \subseteq M$; el complemento $M\setminus M_E$ también es un conjunto. (II) Todos los subconjuntos de un conjunto $M$ forman un conjunto $\mathcal{P}M$. Por cierto, la matemática conjuntista había hecho uso de estos supuestos sin formularlos. Como veremos en la Sección 1.8.1, ambos figuran en la lista de axiomas para la teoría de conjuntos que Zermelo publicó ese mismo año (1908a). El Teorema del Buen Orden se enuncia como una proposición hipotética: Un conjunto $M$ es bien ordenable si hay un selector para $\mathcal{P}M\setminus\{\varnothing\}$. Después de completar la prueba de la misma, Zermelo enuncia el Axioma de Selección, en virtud del cual es lícito desglosar la apódosis: Todo conjunto es bien ordenable. La argumentación de Zermelo se dirige a probar que si $M$ es un conjunto no vacío y $f$ es un selector para $\mathcal{P}M\setminus\{\varnothing\}$, hay un y sólo un $\mathsf{M} \subseteq \mathcal{P}M$ tal que (i) $\mathsf{M}$ está bien ordenado por la relación de inclusión propia '$\mathrm{X} \subseteq \mathrm{Y} \wedge \mathrm{X} \neq \mathrm{Y}$' ('Y incluye a X, pero no es igual a X'), que simbolizaré '$\mathrm{Y} \supset \mathrm{X}$', y (ii) la restricción $f_{\mathsf{M}}$ de $f$ a $\mathsf{M}\setminus\{\varnothing\}$ es una biyección de $\mathsf{M}\setminus\{\varnothing\}$ sobre $M$. Como es obvio, $M$ está bien ordenado entonces por la relación < definida como sigue, para cualquier par de elementos $x, y \in M$: $x < y \Leftrightarrow f_{\mathsf{M}}^{-1}(x) \supset f_{\mathsf{M}}^{-1}(y)$. (Decimos que de esta manera $f_{\mathsf{M}}$ *induce* en $M$ el buen orden de $\mathsf{M}$).[3]

Zermelo responde a objeciones formuladas por Borel y Peano, Poincaré, Philip Jourdain y Schoenflies. Me limitaré a considerar las primeras. Borel y Peano habían reclamado una prueba del Axioma de Selección. Zermelo responde que en matemáticas la indemostrabilidad no equivale a la invalidez, pues, como es sabido, no todo se puede demostrar. Concede que "no puede forzar a nadie apodícticamente a reconocer" el Axioma, pero a la vez arguye que éste reúne los tres requisitos que, a su modo de ver, justifican la adopción de un postulado en matemáticas: (*a*) con frecuencia ha sido *utili-*

[3] En el Apéndice V doy una paráfrasis de la segunda demostración del Teorema del Buen Orden por Zermelo.

*zado* tácitamente en diversos campos de la matemática y especialmente en teoría de conjuntos, (*b*) es *evidente* de suyo y (*c*) responde a una *necesidad científica*, pues son muchas las proposiciones importantes que sólo pueden demostrarse invocándolo. De estas tres condiciones sólo la tercera ha demostrado tener verdadera fuerza. A la lista ofrecida por Zermelo de siete teoremas importantes que no se pueden demostrar sin ayuda del Axioma, han venido a agregarse desde 1908 muchísimos más, al punto de que hoy por hoy el típico especialista en álgebra, topología o análisis diría que si lo privan del Axioma le están quitando el suelo bajo sus pies. Zermelo corona su respuesta con un devastador *argumentum ad hominem*: Peano no necesita el Axioma de Selección en su trabajo profesional debido a que el sistema lógico-matemático expuesto en su *Formulaire des Mathématiques* genera la paradoja de Russell, de la cual se puede deducir lo que a uno le dé la gana.

Más significativa que la objeción de Peano me parece la de Poincaré. En una serie de tres artículos sobre "Las matemáticas y la lógica" (1905/1906), el gran matemático y filósofo francés descargó el peso de su autoridad y el fuego de su elocuencia a la vez sobre el conjuntismo de Cantor y sobre el logicismo de Peano, Russell y Couturat. Poincaré estima que estos programas de fundamentación de las matemáticas andan completamente extraviados. De ahí las paradojas que, a su modo de ver, aquejan a ambos por igual. Según Poincaré, las paradojas nacen del empleo de términos cuya definición envuelve una forma de circularidad qué él juzga viciosa. Así, el término 'conjunto de todos los conjuntos que no son elementos de sí mismos' supuestamente denota un objeto caracterizado mediante una alusión a la totalidad de los conjuntos, uno de los cuales es precisamente ese objeto. Asímismo el ordinal de todos los ordinales, nombrado en la paradoja de Burali-Forti, se define por la expresión 'todos los ordinales', cuya extensión lo contiene. Caracterizar el término $t$ nombrando un determinado conjunto $K$ tal que $t \in K$ es como repetir el *definiendum* en el *definiens*, puesto que cualquier expresión que nombre a $K = \{x\colon x \in K\}$ denota, entre otros, al objeto que se busca designar con $t$. Poincaré objetó a la primera prueba del Teorema del Buen Orden por Zermelo (1904) que la caracterización de los $\gamma$–conjuntos, que figuran decisivamente en ella, peca de este vicio. Zermelo (1908a) no sólo reconoce que ello es así, sino que para explicar a sus lectores la objeción de Poincaré, cita un ejemplo tomado de su segunda prueba, a saber, el término $\mathsf{M}$ que designa la intersección de todas las $f$-cadenas de $M$, la cual —como

luego se prueba— es ella también una de esas $f$-cadenas (v. Apéndice V).

Debido a un desplazamiento semántico que ya se manifiesta en este texto de Zermelo, las definiciones que Poincaré juzga circulares y los términos definidos por ellas suelen distinguirse con el epíteto 'no predicativo' o 'impredicativo'.[4] Zermelo observa que el uso de términos no predicativos es endémico en el análisis: ellos figuran en cada demostración "en que el máximo o el mínimo de un conjunto numérico 'cerrado' definido previamente se utiliza para llegar a nuevas conclusiones. Así ocurre, por ejemplo, en la conocida prueba del Teorema Fundamental del Álgebra por Cauchy, sin que hasta ahora se le haya ocurrido a nadie hallar en ella algo ilógico" (Zermelo 1908, p. 524).[5] Y la verdad es que nadie reconocería un procedimiento falaz en la descripción de *diciembre* como "el último mes del año" o del *perihelio de Mercurio* como "el punto de la órbita de Mercurio que está más cerca del sol". En un diccionario filosófico reciente, Christian Thiel, que milita entre los enemigos de la 'impredicatividad', amaña su definición para evitar los contraejemplos de este género. *Impredicativo*, según Thiel, es "un

4 Russell (1906) llama 'predicativa' a una condición ("función proposicional") que determina un conjunto; 'no predicativa', entonces, es una que no logra hacerlo, como la condición '$x$ es el conjunto de todos los conjuntos'. Poincaré (1905/1906) dice que "las definiciones que deben ser consideradas como no predicativas son aquellas que contienen un círculo vicioso" (CM, p. 147). Zermelo, que probablemente no había leído el artículo de Russell, da la impresión de entender que Poincaré en este pasaje —destacado en cursiva— está *definiendo* el término 'definición no predicativa'. Poincaré acepta aparentemente esta interpretación en su "Logique de l'infini" (1909), cuando distingue "dos especies de clasificaciones aplicables a los elementos de las colecciones infinitas: las clasificaciones *predicativas*, que no pueden ser quebrantadas por la introducción de nuevos elementos, y las clasificaciones *no predicativas*, que la introducción de nuevos elementos obliga a modificar incesantemente" (DP, p. 10). Doy más detalles sobre este asunto en el Capitulo 2.4.

5 El Teorema Fundamental del Álgebra dice que todo polinomio de grado $n \geq 1$, con coeficientes complejos, $a_0 + a_1x^1 + \ldots + a_nx^n$, tiene por lo menos una raíz (en el cuerpo $\mathbb{C}$ de los complejos). Un caso ejemplar de impredicatividad es la definición habitual del *supremo* o cota superior mínima de un conjunto $K \subset \mathbb{R}$: para todo $x$, $z$, y $w \in \mathbb{R}$, $x$ es el *supremo* de $K$ si y sólo si (i) si $w \in K$, entonces $w \leq x$, y (ii) si para cualquier $w \in K$, $z \geq w$, entonces $z \geq x$. (El *ínfimo* o cota inferior máxima de $K \subset \mathbb{R}$ se define análogamente, *mutatis mutandis*). Buena parte del análisis clásico depende de un teorema que dice que todo conjunto de números reales que tiene una cota superior posee un supremo y todo conjunto de números reales que tiene una cota inferior posee un ínfimo.

procedimiento para delimitar o caracterizar un objeto, que en la descripción del mismo hace referencia a una totalidad de objetos que . . . comprendería al propio objeto en cuestión, y *cuyos elementos no pueden todos generarse constructivamente*" (cursiva mía).[6] Conforme a esta nueva definición, claro está, el vicio de impredicatividad no consiste en que se aduzca "circularmente", para fijar la referencia a cierto objeto, una totalidad que lo presupone, sino más bien en que la totalidad en cuestión no satisface un requisito de construibilidad que habría que especificar y justificar. Como la matemática conjuntista no se deja imponer tales requisitos, el desacuerdo entre Poincaré y Zermelo nos sitúa, de hecho, en la línea divisoria entre dos grandes vertientes del pensamiento matemático del siglo XX.

6 Mittlestraß, EPW, s.v. *imprädikativ/Imprädikativität*.

## 1.8 AXIOMAS PARA UNA TEORÍA DE CONJUNTOS

### *1.8.1 Zermelo (1908)*

Respondiendo a la objeción de Peano contra su primera demostración del Teorema del Buen Orden, Zermelo observa que el sistema lógico-matemático de ese autor cae presa de la paradoja de Russell porque no distingue entre 'conjunto' y 'clase'.

> En cambio, como pronto mostraré en otro lugar, los partidarios de la teoría de los conjuntos como disciplina puramente matemática, que no está constreñida por los conceptos fundamentales de la lógica tradicional, son perfectamente capaces de evitar, mediante una apropiada especialización de sus axiomas, todas las "antinomias" conocidas hasta ahora.
>
> (Zermelo 1908, p. 115)

El escrito anunciado apareció ese mismo año (Zermelo 1908a). Con él nace la teoría axiomática de conjuntos.

La idea de una teoría axiomática procede de Aristóteles, para quien todo conocimiento científico propiamente tal (ἐπιστήμη) se establece por inferencia deductiva a partir de principios (ἀρχαί) de dos clases, a saber, conceptos que no se definen y aseveraciones que no se demuestran. En la literatura filosófica posterior se los llama, respectivamente, 'primitivos' y 'axiomas'. Los primitivos no tienen que definirse porque cualquiera los entiende. No es posible, pero tampoco es preciso, demostrar los axiomas, porque son de suyo evidentes.[1] Tradicionalmente se ha visto en los *Elementos* de Euclides (publicados alrededor de un cuarto de siglo después de muerto Aristóteles) una realización ejemplar de esta idea de ciencia. No comparto esta opinión.[2]

[1] Aristóteles explica su idea de una ciencia deductiva en los *Segundos Analíticos*. Scholz (1930) ofrece una interpretación de esa obra a la luz de la axiomática moderna (en el estilo de Hilbert). Sobre esta materia, puede también consultarse mi artículo, "El método axiomático" (1993).

[2] Véase Torretti 1978, pp. 5–9.

A mi modo de ver, la concepción aristotélica de una ciencia edificada por deducción rigurosa sobre principios de suyo inteligibles o evidentes no vino a realizarse hasta 1882, en las *Lecciones sobre geometría moderna* de Moritz Pasch.

La exposición sistemática de la geometría en el libro de Pasch parte de ciertos conceptos primitivos concernientes a la figura, tamaño y posición recíproca de los cuerpos. Dichos conceptos no se definen, pues ninguna definición podría sustituir a "la indicación de objetos naturales apropiados", que es la única vía para entender nociones tan simples e irreductibles como éstas (Pasch 1882, p. 16). Todos los demás conceptos utilizados se definen en términos de los primitivos o de otros conceptos ya definidos. Los axiomas (*Grundsätze*) conectan los conceptos primitivos entre sí. Aseveran "lo que se observa en ciertas figuras muy simples" y su contenido geométrico no puede captarse aparte de esas figuras (1882, p. 43). Todas las demás aseveraciones de la geometría —sus "teoremas" (*Lehrsätze*)— se deducen de los axiomas o de teoremas ya deducidos. Así, "todo lo que se necesita para probar los teoremas tiene que estar consignado (*niedergelegt*), sin excepción, en los axiomas" (1882, p. 5). "Cada conclusión a la que se llega en una prueba tiene que hallar su confirmación en la figura, pero no se justifica por la figura sino por una determinada proposición (o definición) precedente" (1882, p. 43).

En agudo contraste con el mismo Aristóteles y la larga tradición de sus intérpretes, Pasch comprendió a cabalidad las implicaciones de esta idea de una ciencia rigurosamente deductiva.

> De hecho, si la geometría ha de ser realmente deductiva, el proceso de inferencia debe ser siempre independiente del *sentido* de los conceptos geométricos, tal como debe ser independiente de las figuras. Sólo pueden considerarse las *relaciones* entre los conceptos geométricos consignadas en las proposiciones y definiciones utilizadas. Durante la deducción es sin duda lícito y útil pensar en el significado de los conceptos geométricos que se presentan, pero ello *no es en absoluto necesario*; de modo que cuando resulta indispensable, esto indica, justamente, que la deducción tiene lagunas y —si no es posible llenarlas modificando el razonamiento— que las proposiciones aducidas como medios de prueba son insuficientes.
>
> (Pasch 1882, p. 98)

Este pasaje anticipa el nuevo giro que Hilbert le dará poco más tarde a la idea de una ciencia deductiva, y que inspirará, prácticamente sin rivales, todas

las axiomatizaciones de teorías matemáticas en el siglo XX. Si el significado de los conceptos primitivos —fijado intuitivamente por "indicación de objetos naturales apropiados"— nada contribuye a la justificación de los teoremas, la teoría deductiva en cuestión valdrá lo mismo para el sistema de objetos indicado que para cualquier otro sistema de objetos cuyas relaciones mutuas satisfagan las condiciones estipuladas en los axiomas. Así, por ejemplo, la teoría del espacio euclidiano axiomatizada por Hilbert retiene todo su valor si se entiende que un "punto" es un trío ordenado de números reales, que una "recta" o un "plano" es un cierto conjunto de "puntos" que satisfacen un cierto tipo de ecuaciones, se interpretan los otros primitivos de Hilbert de un modo apropiado —por ejemplo, el "punto" $\mathbf{x} = \langle x_1,x_2,x_3\rangle$ "incide" en el plano $\pi$, si los números $x_1$, $x_2$ y $x_3$ forman una solución del sistema de ecuaciones lineales constitutivo de $\pi$— y se define la "distancia" $\delta(\mathbf{v},\mathbf{u})$ entre dos "puntos" $\mathbf{v} = \langle v_1,v_2,v_3\rangle$ y $\mathbf{u} = \langle u_1,u_2,u_3\rangle$ por la relación pitagórica

$$\delta(\mathbf{v},\mathbf{u}) = \left(\sum_{i=1}^{3}\left(v_i - u_i\right)^2\right)^{1/2}$$

Como Hilbert le explica a Frege el 29 de diciembre de 1899:

> Naturalmente, cada teoría es sólo un andamiaje (*Fachwerk*) o esquema de conceptos con sus necesarias relaciones mutuas, y los elementos básicos pueden pensarse como se quiera (*in beliebiger Weise*). Si pienso que mis puntos son cualquier sistema de cosas, vgr. el sistema *amor*, *ley*, *deshollinador*, . . . , con que luego sólo postule la totalidad de mis axiomas como relaciones entre estas cosas, mis teoremas —el de Pitágoras, por ejemplo— valen también para ellas. En otras palabras: cada teoría puede siempre aplicarse a infinitos sistemas de elementos básicos. Basta aplicar una transformación unívoca invertible [una biyección—R.T.] y estipular que los axiomas homólogos valen para las transformadas.
>
> [Hilbert a Frege, 29.12.1899, en Frege, WB, p. 67]

Las lúcidas observaciones de Pasch y Hilbert que he citado dan al traste con la idea de Gergonne (1818), repetida por varios autores a comienzos de este siglo, según la cual los axiomas de una teoría proveen una *definición implícita* de sus primitivos. De hecho, es muy poco o nada lo que los axiomas pueden decir acerca de la índole de los objetos que los primitivos denotan. Es claro, en cambio, que los axiomas definen *explícitamente* una clase de

objetos complejos, a la que pertenece cada sistema de elementos básicos que cumple con las condiciones estipuladas por ellos. Si a una lista de axiomas compatibles entre sí viene a agregarse otro que no es una consecuencia de los anteriores, la clase definida se restringe. (En particular, si el nuevo axioma es incompatible con los demás, la clase definida se reduce a $\varnothing$). Cuando se habla de "especializar" los axiomas de una disciplina matemática, uno pensaría normalmente que se trata de restringir de este modo el alcance de sus aplicaciones potenciales. Por ejemplo, si eliminamos el Axioma IV (Axioma de las Paralelas) de la lista propuesta en *Los fundamentos de la geometría* de Hilbert (1899), la clase de objetos complejos caracterizada por la teoría resultante incluye todos los ejemplos concebibles de espacio euclidiano y todos los ejemplos concebibles de espacio de Lobachevsky (entiéndase: de tres dimensiones). Al agregar el Axioma IV la teoría se especializa y deja de ser aplicable a los espacios de Lobachevsky.[3] En cambio, la teoría obtenida agregando la negación del Axioma IV caracteriza precisamente a los espacios de Lobachevsky y no es aplicable a los euclidianos. Por último, si agregamos a los Axiomas I–III y V de Hilbert tanto el Axioma IV como su negación, obtenemos una teoría tan especializada que no podría haber nada a lo que fuese aplicable. Sorprende, pues, que Zermelo pretendiese limpiar de contradicciones a la teoría de conjuntos por la vía de *especializar* sus axiomas.

Aunque influida por la obra de Hilbert, la teoría axiomática de conjuntos de Zermelo tiene características peculiares, impuestas por la naturaleza de su tema. Una que salta a la vista concierne a la manera de referirse a éste. Normalmente, una teoría axiomática se refiere a uno o más *conjuntos* de objetos, cuyos atributos y relaciones son caracterizados por los axiomas.[4] Pero como la teoría de Zermelo intenta justamente caracterizar el atributo de *ser un conjunto* no puede invocarlo de entrada al acotar su tema. Leemos, por eso, que "la teoría de conjuntos concierne a un '*dominio*' $\mathfrak{B}$ de objetos

[3] El Axioma IV de Hilbert dice que si $a$ es una recta cualquiera y $A$ es un punto fuera de $a$, en el plano determinado por $a$ y $A$ hay a lo sumo una recta que pasa por $a$ y no intersecta a $A$. El Axioma IV es falso en un espacio de Lobachevsky. En tal espacio no hay rectángulos y el cuarto ángulo de un cuadrilátero con tres ángulos rectos es siempre agudo.

[4] Hilbert invita al lector a concebir "tres sistemas de cosas" (que llama 'puntos', 'rectas' y 'planos'), donde —como señalé en la nota 3 del Capitulo 1.2— 'sistema' (*System*) es el término que usaba Dedekind en lugar de 'conjunto' (*Menge*).

que simplemente llamamos ‘*cosas*’, una parte de las cuales está formada por los ‘conjuntos’.”[5] Para comprender esta oración no tenemos que saber lo que significa la palabra ‘conjunto’, ya que esto es lo que los axiomas buscan determinar, pero ciertamente tenemos que entender la frase ‘dominio de objetos’. Ahora bien, se preguntará el lector ¿no denota esta frase exactamente lo que hasta aquí, siguiendo a Cantor, hemos llamado ‘conjunto’? ¿asistimos, entonces, a un acto de prestidigitación verbal? En la Sección 1.8.5 presentaré la profunda y original respuesta a tales preguntas que años más tarde ofrecerá Zermelo (1930). Pero por ahora podemos encarar este asunto así: la expresión ‘el conjunto $K$’, en su acepción cantoriana, designa *colectivamente* a los elementos de $K$; en cambio, la frase ‘el dominio $\mathfrak{B}$’ designa *distributivamente* a las cosas cuyos atributos y relaciones Zermelo busca caracterizar con sus axiomas. En efecto, si ‘el dominio $\mathfrak{B}$’ designase un objeto formado por esas cosas, se podría discernir en $\mathfrak{B}$ otro objeto $\mathfrak{C}$, formado por aquellas cosas de $\mathfrak{B}$ que son conjuntos; entonces $\mathfrak{C}$, no importa como se lo describa, sería en efecto el *conjunto de todos los conjuntos*, cuya existencia es contradictoria.

Otra diferencia notoria entre la teoría de Zermelo y las axiomáticas habituales concierne al número y complejidad de los primitivos. Mientras la geometría de Hilbert tiene ocho, a saber, tres predicados monádicos (‘punto’, ‘recta’, ‘plano’), cuatro diádicos (dos especies de incidencia y dos de congruencia) y uno triádico (‘el punto $x$ está entre los puntos $z$ y $w$’), la teoría de Zermelo sólo tiene un predicado monádico, ‘$x$ es un conjunto’, y uno diádico, ‘$x$ es un elemento del conjunto $y$’ (simbolizado ‘$x \in y$’). Entre esos ocho primitivos, los axiomas de Hilbert postulan relaciones sumamente complejas, las cuales caracterizan su dominio con tal precisión que, a fin de cuentas, cualquier modelo (realización) de la geometría de Hilbert es isomórfico[†] a cualquier otro. Si una teoría axiomática posee esta propiedad se dice que es *monomórfica* o *categórica*. La teoría de Zermelo no es categórica,

5 “Die Mengenlehre hat zu tun mit einem ‘*Bereich*’ $\mathfrak{B}$ von Objekten, die wir einfach als ‘*Dinge*’ bezeichnen wollen, unter denen die ‘Mengen’ einen Teil bilden” (Zermelo 1908a, p. 262). En virtud del Axioma II si $a$ es una “cosa” del dominio $\mathfrak{B}$ hay en $\mathfrak{B}$ por lo menos un conjunto $k$ tal que $a \in k$. Así, pues, todas las “cosas” de que habla la teoría axiomática de Zermelo son elementos de conjuntos. Aquéllas que no son conjuntos a su vez se conocen en la literatura matemática como *Urelemente* (‘elementos primordiales’).

y podría pensarse por ello que ofrece una caracterización insuficiente de su tema. Para muchos éste es un defecto de la teoría axiomática de conjuntos que Fraenkel intentará subsanar pero que según Skolem y von Neumann no tiene remedio. Por su parte, Zermelo (1930) sostendrá que la no-categoricidad es una virtud imprescindible en una teoría que pretende enmarcar la matemática entera.[6]

En el Apéndice VI doy una traducción de los axiomas de Zermelo. Aquí comentaré su alcance, en parte citándolos, en parte parafraseándolos. Ante todo, quiero llamar la atención sobre un hecho notable: salvo el Axioma I, todos los axiomas de Zermelo son asertos de existencia, ya sea absolutos, ya sea condicionales: los Axiomas II y VII aseveran que en el dominio $\mathfrak{B}$ hay un conjunto de cierta índole; los Axiomas II–VI aseveran que, si en $\mathfrak{B}$ hay ciertos objetos o conjuntos, también hay en $\mathfrak{B}$ ciertos otros conjuntos que cumplen ciertas condiciones con respecto a aquéllos. Los asertos de existencia son frecuentes en todas las teorías axiomáticas, pero éstas incluyen también lo que —por analogía con la física— cabe llamar "leyes", esto es, axiomas que postulan que si unos objetos del dominio cumplen una condición, también cumplen otra.[7] Parecería que sin tales "leyes" no se puede especificar la naturaleza del dominio de la teoría. La única "ley" entre los axiomas de Zermelo es el Axioma I (Axioma de Determinación), en virtud del cual la identidad de un conjunto sólo depende de sus elementos. También tiene este carácter la indicación preliminar de que, con la sola excepción expresada en el Axioma II, una cosa *a* es un conjunto si y sólo

6 "Nuestro sistema de axiomas justamente es *no-categórico*, lo que en este caso no es un inconveniente, sino una *ventaja*. Pues precisamente sobre este hecho descansa la enorme significación y la aplicabilidad ilimitada de la teoría de conjuntos en general" (Zermelo 1930, p. 45). Zermelo se refiere aquí a su sistema *ZF′* de 1930, pero el pasaje también es aplicable al sistema de 1908a.

7 El Axioma I 8 de Hilbert es un buen ejemplo de aserto existencial absoluto: "Hay por lo menos cuatro puntos que no están todos en un mismo plano". El Axioma II 2 es un aserto existencial condicional: "Dados dos puntos *A* y *C*, hay siempre un punto *B* sobre la recta *AC* tal que *C* está entre *A* y *B*." Entre los 18 primeros axiomas de Hilbert sólo siete no son asertos de existencia: los Axiomas I 5 y II 3 niegan la existencia de ciertos objetos; los Axiomas I 6, II 1, III 2, III 3 y III 5 afirman que ciertas relaciones subsisten cada vez que se cumplen ciertas relaciones. Los dos axiomas restantes son peculiares. El Axioma V afirma que, dados dos segmentos rectos $\alpha$ y $\beta$, tales que $\alpha < \beta$, siempre hay un número entero $n$ tal que $\beta < n\alpha$. Se trata pues de un aserto existencial, pero, como $n$ no es un punto, una recta, o un plano, lo que V 1 asevera

si hay una cosa $b$ tal que $b \in a$ (Zermelo 1908, p. 262, §1.2). Eso es todo lo que la teoría nos brinda para visualizar qué podría ser un 'conjunto' y en qué consiste la relación '$\in$'.

Los asertos existenciales absolutos establecen el contenido mínimo del dominio $\mathfrak{B}$.

Conforme al Axioma II (Axioma de los Conjuntos Elementales), hay un conjunto tal que no contiene elementos, "el conjunto cero", simbolizado "0". Me parece justificado identificarlo sin más con el conjunto vacío $\varnothing$. Obsérvese que lo que aquí se está diciendo es (*a*) que hay en $\mathfrak{B}$ (por lo menos) una cosa, (*b*) que (por lo menos) una de las cosas que hay en $\mathfrak{B}$ no contiene elementos, y (*c*) que no obstante la característica general atribuida a los conjuntos por la indicación preliminar (§ 1.2), cierta cosa $\varnothing$ que no contiene elementos será tratada como conjunto en nuestro discurso. Evidentemente, como la identidad de un conjunto depende de sus elementos (Axioma I), este privilegio no puede conferirse más que a una sola cosa. Completan el Axioma II dos aseveraciones condicionales: si hay en $\mathfrak{B}$ una cosa cualquiera $a$, entonces también hay en $\mathfrak{B}$ un conjunto $\{a\}$, cuyo único elemento es $a$; si hay en $\mathfrak{B}$ dos cosas $a$ y $b$, entonces también hay en $\mathfrak{B}$ un conjunto $\{a,b\}$, cuyo únicos elementos son $a$ y $b$.

Conforme al Axioma VII (Axioma del Infinito), hay en $\mathfrak{B}$ por lo menos un conjunto $Z$ tal que (i) $\varnothing \in Z$ y (ii) si $a \in Z$, $\{a\} \in Z$. Invocando el Axioma III, Zermelo demuestra que existe en $\mathfrak{B}$ un cierto conjunto $Z_0$ —que es la parte común a todos los conjuntos $Z$ que cumplen la condición del Axioma VII— cuyos elementos son $\varnothing$, $\{\varnothing\}$, $\{\{\varnothing\}\}, \ldots$ Zermelo propone llamar a $Z_0$ "la serie numérica", porque sus elementos puede hacer las veces de números. "Constituye el ejemplo más simple de un conjunto 'enumerablemente infinito'" (Zermelo 1908, p. 267).

Los asertos existenciales condicionales postulan una expansión colosal del contenido de $\mathfrak{B}$.

Como vimos, si hay dos cosas distintas, $a$ y $b$, las cláusulas condicionales del Axioma II certifican la existencia de los infinitos conjuntos $\{a\}$, $\{\{a\}\}$, $\{\{\{a\}\}\}, \ldots, \{b\}$, $\{\{b\}\}$, $\{\{\{b\}\}\}, \ldots$, $\{a,b\}$, $\{a,\{b\}\}$, $\{\{a\},\{a,b\}\}$, etc.

---

es la existencia de un objeto ajeno al dominio de la teoría. También el Axioma V 2 (trascrito más adelante en la nota 27) trasciende el dominio de la teoría, pues dice que fuera de él no hay otros objetos con los cuales se lo pueda ampliar para constituir un dominio más rico en que también se cumplan los axiomas.

En virtud del Axioma III (Axioma de Separación), los elementos de un conjunto *a* existente en $\mathfrak{B}$ que cumplen una condición bien definida *P* forman un conjunto separado existente en $\mathfrak{B}$, el subconjunto $\{x: x \in a \wedge Px\}$. Para que este axioma tenga un significado preciso hay que determinar qué es una condición bien definida (*definit*). Zermelo lo explica así:

> Una cuestión o proposición $\mathfrak{E}$, cuya validez o invalidez deciden sin arbitrariedad las relaciones fundamentales del dominio [esto es, las relaciones de la forma $a \in b$] en virtud de los axiomas y de las leyes lógicas universalmente válidas se llama "*bien definida*" ["*definit*"]. Asímismo, una proposición universal $\mathfrak{E}(x)$, en la cual el término variable *x* puede recorrer todos los individuos de una clase $\mathfrak{K}$, se dice "bien definida" si está bien definida para *cada* individuo de la clase $\mathfrak{K}$. Así, la cuestión de si $a \in b$ o no está siempre bien definida; también la cuestión de si $M \subseteq N$ o no.
> (Zermelo 1908a, p. 263)

Aunque a primera vista parece que la frase "deciden sin arbitrariedad" (*ohne Willkür entscheiden*) no es suficientemente precisa (Weyl 1910, p. 304), una breve reflexión permite extraer del pasaje citado la siguiente definición recursiva: Una condición *P* está bien definida (*a*) si *P* expresa la presencia o ausencia de la relación $\in$ entre dos cosas; (*b*) si *P* se refiere a *todas* las cosas que cumplen una cierta condición bien definida con respecto a cada una de ellas; o (*c*) si *P* se infiere de otras condiciones bien definidas, en virtud de los Axiomas I–VII y de las leyes de la lógica.[8] Cabe todavía preguntarse qué hay que entender aquí por 'leyes de la lógica'. Como veremos en la Sección 1.8.2, Skolem (1922) dará a esta pregunta una respuesta radical que ha tenido mucha aceptación.

El Axioma IV (Axioma del Conjunto Potencia) dice que si *T* es un conjunto existente en $\mathfrak{B}$, también existe en $\mathfrak{B}$ el *conjunto potencia* $\mathcal{P}T$ cuyos elementos son todos los subconjuntos de *T*. Combinado con el Axioma VII, que postula la existencia en $\mathfrak{B}$ de un conjunto infinito (enumerable), el Axioma IV asegura la existencia en $\mathfrak{B}$ de conjuntos indenumerables cuyos

[8] Apliquemos esta definición a los ejemplos de 'cuestión bien definida' propuestos por Zermelo. '$a \in b$' está siempre bien definida en virtud de (*a*). '$M \subseteq N$' abrevia la condición $\forall x(x \in \mathrm{M} \supset x \in \mathrm{N})$, la cual está bien definida en virtud de (*b*), ya que '$x \in \mathrm{N}$' —en virtud de (*a*)— y por ende '$x \in \mathrm{M} \supset x \in \mathrm{N}$' —en virtud de (*c*)— son condiciones bien definidas para toda cosa *x*.

cardinales son $2^{\aleph_0}$, $2^{2^{\aleph_0}}$, . . .

El Axioma V (Axioma de Unión) dice que si $T$ es un conjunto existente en $\mathfrak{B}$, también existe en $\mathfrak{B}$ el *conjunto unión* $\bigcup T$ formado por todos los elementos de los elementos de $T$.

El Axioma VI es el Axioma de Selección. Formalmente: si $T$ es un conjunto existente en $\mathfrak{B}$ cuyos elementos son conjuntos no vacíos y mutuamente disjuntos (esto es, si $u \in T$ implica que $\exists w(w \in u)$ y $u \in T \wedge v \in T$ implica que $u = v \vee u \cap v = \varnothing$), el conjunto unión $\bigcup T$ incluye un subconjunto $s$ que tiene un y sólo un elemento en común con cada elemento de $T$.[9] Informalmente, esto equivale a decir que si $T$ cumple la condición indicada, es posible seleccionar un y sólo un elemento de cada elemento de $T$, y formar con los objetos así seleccionados un conjunto $s$.

De los siete axiomas citados Zermelo infiere los principales teoremas de la teoría de los cardinales transfinitos (1908a, § 2 "Teoría de la equivalencia"). La teoría de los ordinales, cuya publicación anuncia (p. 262), y que se sabe que completó hacia 1914, no llegó a publicarse.

Se ha solido decir que Zermelo elaboró su teoría axiomática de conjuntos con el propósito de eludir las paradojas. G. H. Moore sostiene que esta opinión tradicional contiene a lo sumo "una migaja de verdad" (1982, p. 157), pero que Zermelo obró motivado sobre todo por el deseo de responder a los críticos de su primera prueba del Teorema del Buen Orden, para lo cual necesitaba determinar con toda precisión los supuestos requeridos para demostrarlo. Como no he tenido acceso a la correspondencia u otros papeles privados de Zermelo, no puedo cuestionar este aserto de Moore. Lo que no admite duda es que Zermelo *ofreció al público* su axiomatización como el método apropiado para reconstruir la teoría de Cantor sin paradojas. Lo dice el anuncio de Zermelo 1908a en Zermelo 1908 (citado al comienzo de esta sección) y lo confirma ampliamente la introducción de Zermelo 1908a. Leemos allí que la "Antinomia de Russell" relativa al "conjunto de todos los conjuntos que no se contienen a sí mismos" nos fuerza a restringir la definición cantoriana de conjunto como una "reunión en un todo de determinados objetos bien distinguidos de nuestra intuición o nuestro pensamiento", aunque "todavía no se ha logrado reemplazarla con otra, igualmente simple,

9 La apódosis del Axioma VI se puede expresar con más claridad así:

$$\exists s(s \subseteq \bigcup T \wedge \forall u(u \in T \supset (u \cap s \neq \varnothing \wedge \forall x \forall y(\{x,y\} \subseteq u \cap s \supset x = y))))$$

que no dé lugar a tales dudas” (1908a, p. 261).

> En estas circunstancias no queda en la actualidad otro recurso que emprender el camino al revés y, partiendo de la “teoría de conjuntos” históricamente dada, buscar los principios que se requieren para fundamentar esta disciplina matemática. Esta tarea debe resolverse de tal modo que los principios se restrinjan lo suficiente para excluir todas las contradicciones, pero a la vez sean lo bastante amplios para retener todo lo que hay de valioso en dicha teoría.
>
> (Zermelo 1908a, p. 261)

Zermelo estima que sus siete axiomas cumplen este propósito. Confiesa que no ha podido aún demostrar rigurosamente su consistencia (*Widerspruchslosigkeit*), pero subraya que “todas las ‘antinomias’ conocidas hasta ahora desaparecen en cuanto se adoptan como base los principios aquí propuestos” (1908a, p. 262).

### *1.8.2 ¿Qué está ‘bien definido’?*

Cuando Zermelo publica sus axiomas en 1908, ya está en marcha el esfuerzo iniciado por Frege y Peano para introducir precisión y claridad en el pensamiento matemático reformulándolo en un lenguaje artificial de sintaxis austera y rigurosa. Esta empresa intelectual, a la que me refiero en la Segunda Parte, tuvo una marcada influencia en el desarrollo ulterior de la teoría axiomática de conjuntos. Ello no obstante, y a riesgo de adelantar ideas que debo explicar allí, me ha parecido oportuno cerrar esta Primera Parte con una descripción sumaria de las principales innovaciones que median entre el primer sistema axiomático de Zermelo (1908a) y los sistemas ZF (Zermelo-Fraenkel) y NGB (von Neumann-Gödel-Bernays) utilizados corrientemente en la enseñanza y la investigación.

En la conferencia “Sobre las definiciones de los conceptos matemáticos fundamentales”, dictada en Zurich al obtener la *venia legendi*, Weyl (1910) expresa reservas frente a la noción de lo ‘bien definido’ (*definit*) invocada en el Axioma III de Zermelo, cuya caracterización parafrasea así:

> Una proposición bien definida es una proposición tal que su verdad o falsedad se puede decidir únivocamente y sin arbitrariedad sobre la base de las relaciones fundamentales $\in$ que subsisten entre las cosas de la teoría de conjuntos.
>
> (Weyl 1910, p. 304)

Weyl observa con razón que hablar de una "decisión unívoca y exenta de arbitrariedad" es un tanto vago, y propone en cambio una definición recursiva más precisa pero posiblemente más estrecha que la que extraje del texto original de Zermelo en la Sección 1.8.1. De hecho, Weyl no llega a formular tal definición, sino que nos invita a construirla por analogía con ciertos "principios de definición" de los conceptos geométricos enunciados en la misma conferencia (1910, pp. 299s.). Es lo que intento a continuación. (El importante concepto de *definición recursiva* se explica en el Apéndice VIII. En el siguiente ejemplo, la cláusula señalada con el índice 0 constituye la *base* de la recursión):

(0) Toda proposición de la forma '$x \in y$' está bien definida, cualesquiera que sean los términos $x$ e $y$.

(1) Si $\varphi$ está bien definida, también lo está su negación $\neg\varphi$.

(2) Si $\varphi$ y $\psi$ están bien definidas, también lo está su conjunción $\varphi \wedge \psi$.

(3) Si $\varphi(\alpha_1, \dots, \alpha_\nu)$ es una proposición bien definida, referente al $\nu$-tuplo de cosas $\langle\alpha_1, \dots, \alpha_\nu\rangle$, y $\sigma$ designa una permutación de $\{1, \dots, \nu\}$, entonces la proposición $\varphi(\alpha_{\sigma 1}, \dots, \alpha_{\sigma\nu})$, referente al $\nu$-tuplo $\langle\alpha_{\sigma 1}, \dots, \alpha_{\sigma\nu}\rangle$, también está bien definida.

(4) Si $\varphi(\alpha_1, \dots, \alpha_\nu)$ está bien definida, y $\forall\xi\psi(\alpha_1, \dots, \alpha_\nu, \xi)$ es verdadera si y sólo si $\varphi(\alpha_1, \dots, \alpha_\nu)$ es verdadera, entonces $\psi(\alpha_1, \dots, \alpha_\nu, \alpha_{\nu+1})$ está bien definida cualquiera que sea la cosa designada por $\alpha_{\nu+1}$.

(5) Si $\varphi(\xi_1, \dots, \xi_\nu)$ está bien definida, y $\psi(\alpha_1, \dots, \alpha_{\nu-1})$ es verdadera si y sólo si $\exists\xi\varphi(\alpha_1, \dots, \alpha_{\nu-1}, \xi)$ es verdadera, $\psi(\alpha_1, \dots, \alpha_{\nu-1})$ está bien definida.

(6) $\varphi$ es una proposición bien definida si y sólo si es posible caracterizarla como tal mediante un número finito de aplicaciones de las seis reglas precedentes.

En 1910 Weyl todavía estaba dispuesto a decir que la matemática, al menos en lo que respecta a su “contenido lógico”, es “la ciencia de $\in$ y de aquellas relaciones que se dejan definir sobre la base de este concepto conforme a los principios antedichos” (1910, p. 304). Advierte, con todo, una dificultad para esta manera de ver en el hecho siguiente: al hablar de “un número finito de aplicaciones” de ciertos “principios de definición”, tratamos a dichos principios “como cosas enumerables en el sentido de la teoría de conjuntos”; así, estamos presuponiendo esta teoría en la explicación misma de los axiomas en que pretendemos basarla. En 1917, en su monografía sobre *El continuo*, reconoce que no ha podido resolver esta dificultad y cambia de punto de vista:

> El intento de adoptar estos principios [de definición] como axiomas para la formación de conjuntos y de expresar la exigencia de que no existan más conjuntos que los que pueden formarse mediante la aplicación, un número finito de veces, de los principios de construcción contenidos en los axiomas, todo ello *sin presuponer el concepto de los números naturales*, me condujo a una vasta y cada vez más compleja formalización, sin que lograra un resultado definitivo. Sólo en conexión con ciertos conocimientos filosóficos generales [ . . . ] se me hizo claro que perseguía un pseudo-problema escolástico, y alcancé el firme convencimiento [ . . . ] de que la *representación de la iteración, de la serie natural de los números, es un fundamento último del pensamiento matemático*. [ . . . ] Si es verdad que los conceptos fundamentales de la teoría de conjuntos sólo pueden captarse llevando a cabo esta intuición “pura”, es superfluo y confundidor basar a su vez el concepto de número natural en la teoría de conjuntos.
>
> (Weyl 1918, pp. 36–37)

En *El continuo*, Weyl ensaya una filosofía construccionista a medio camino entre la ortodoxia conjuntista a que adhería en 1910 y el intuicionismo brouweriano que abrazará públicamente en 1921. Esta obra cae fuera de los límites del presente estudio. En cambio, debemos atender a varias formas de definir *definit* surgidas en los años 20 dentro de la corriente central del pensamiento matemático. Aunque todas son variantes de la idea propuesta por Weyl en 1910, los autores que voy a citar deben haberla redescubierto independientemente, puesto que no reconocen su deuda con él.

Fraenkel (1922a) demuestra que el Axioma de Selección es independiente de los otros seis axiomas postulados por Zermelo (1908a).[10] Para ello nece-

sita una formulación precisa del Axioma de Separación, que en el escrito de Zermelo "contiene la imprecisa noción *definit*." Fraenkel se vale en cambio de un concepto de *función* que, parafraseando sus palabras, puede caracterizarse así: Sea $x$ un elemento cualquiera de un dado conjunto $K$; entonces $\varphi(x)$ es el objeto que se forma aplicando un número finito de veces los Axiomas II–VI al objeto "variable" $x$ y posiblemente también a otros objetos dados —"constantes"— $a_1, \ldots, a_n$. La regla que gobierna dicha aplicación de los Axiomas II–VI es la *función* designada con la letra $\varphi$. Por ejemplo, $\varphi(x) = \bigcup\{\{\{x\}, \{\varnothing\}\}, \mathcal{P}x \cup \{\{\varnothing\}\}\}$. Fraenkel subraya que este concepto "no incluye ninguna noción general de función o correspondencia ni otra noción fundamental nueva" (1922a, en Heijenoort 1967, p. 286). Mediante este concepto el Axioma de Separación puede reformularse así:

Axioma III. Si $M$ es un conjunto y $\varphi$ y $\psi$ son dos funciones definidas en ese orden, $M$ incluye (i) un subconjunto $M_{\mathfrak{E}}$ que contiene precisamente todos los elementos $x \in M$ tales que $\varphi(x) \in \psi(x)$ y ningún otro elemento; así como (ii) un subconjunto $M_{\neg\mathfrak{E}}$ que contiene precisamente todos los elementos $x \in M$ tales que $\varphi(x) \notin \psi(x)$ y ningún otro elemento.

Por cierto, en la prueba de la independencia del Axioma VI se sobreentiende que las funciones $\varphi$ y $\psi$ mencionadas en el Axioma III sólo envuelven aplicaciones de los Axiomas II–V.[11]

Zermelo (1929) considera que esta nueva versión de su Axioma III es objetable porque emplea la expresión 'un número finito de veces' en la definición de 'función'. Ve aquí un vicio de circularidad, puesto que el concepto mismo de número debiera fundarse en la teoría axiomática de conjuntos. Zermelo concuerda en esto con la autocrítica de Weyl, mas no por ello depone sus convicciones conjuntistas. Antes bien, mantiene la versión original del Axioma III, suplementándola con lo que describe como una caracterización

10 Una proposición $P$ es *independiente* de un conjunto de proposiciones $\mathcal{S}$ si $P$ no es una consecuencia lógica de $\mathcal{S}$. Como toda proposición es una consecuencia lógica de $\mathcal{S}$ si $\mathcal{S}$ es inconsistente, la tesis de que $P$ es independiente de $\mathcal{S}$ se entiende siempre sujeta a la condición tácita de que $\mathcal{S}$ sea consistente.

11 En el Apéndice VII bosquejo la prueba de que el Axioma VI de Zermelo es independiente de los Axiomas I–V y VII (Fraenkel 1922a).

axiomática del concepto *definit*. De hecho lo que nos da es una definición recursiva, suplementada por un curioso "axioma de completud" calcado sobre el Axioma V 2 de los *Fundamentos de la geometría* de Hilbert.[12]

Antes de enunciar la definición de 'definit', Zermelo prescribe una exigencia que este concepto debe satisfacer. Una lista de axiomas *A* —dice— determina un "sistema lógicamente cerrado", esto es, un sistema *S* de proposiciones que pueden derivarse de *A* de un modo puramente lógico. Si *A* es consistente "tiene también que ser '*realizable*', es decir, representable mediante un '*modelo*', una *matriz completa* de las '*relaciones fundamentales*' que figuran en [*S*]" (Zermelo 1929, p. 341). Ahora bien, si $\mathbf{r}(x_1, \ldots, x_n)$ es una de las relaciones fundamentales que figuran en el sistema *S* y *B* es el dominio de elementos de un tal modelo de *A*, entonces tiene que estar decidido para cualquier lista $\langle a_1, \ldots, a_n \rangle$ de objetos de *B* si vale para ella la relación **r** o bien su negación $\neg\mathbf{r}$.

> Una tal propiedad o relación decidida (*entschiedene*) *en cada modelo* por las relaciones fundamentales es lo que se quiso designar con la expresión "*definit*" en [Zermelo 1908a]. "Definit" es, pues, lo que está *decidido en cada modelo particular*, pero puede estarlo de distinta manera en distintos modelos. La "decisión" se refiere al *modelo* particular; la "definición" ("*Definitheit*") misma a la *relación* considerada y al *sistema* total.
>
> (Zermelo 1929, p. 341)

La definición de 'definit' va precedida de una estipulación terminológica que me parece oscura y parafrasearé según mis luces:

> Sea $\mathfrak{B}$ un dominio (que puede constar de varios subdominios mutuamente exclusivos $\mathfrak{B}_1, \ldots, \mathfrak{B}_m$). Sea $\mathfrak{R}$ un sistema de relaciones fundamentales de la forma $\mathbf{r}(x_1, \ldots, x_n)$, donde las variables $x_1, \ldots,$

[12] En la nota 27 cito el Axioma V 2 de Hilbert. El "axioma de completud" de Zermelo dice así:

> AXIOMA III. Si *P* es el sistema de *todas* las proposiciones "bien definidas" o, más generalmente, si es *cualquier* sistema de proposiciones *p* de la índole *Dp*, *P* no incluye ningún subsistema propio $P_1$ que contenga todas las *relaciones fundamentales* de $\mathfrak{R}$ y además comprenda todas las negaciones, conjunciones, disyunciones y generalizaciones de las proposiciones o funciones proposicionales contenidas en él.
>
> (Zermelo 1929, p. 344)

$x_n$ recorren el dominio $\mathfrak{B}$ (pudiendo también restringirse el recorrido de cada variable $x_k$ a uno solo de los subdominios de $\mathfrak{B}$).[13]

Zermelo propone escribir $D\varphi$ en lugar de la expresión '*la proposición* $\varphi$ *está bien definida con respecto a* $\mathfrak{R}$' que define recursivamente así:

0. Si $\mathbf{r}(x_1, \ldots, x_n)$ es una expresión en la cual **r** designa una relación $n$-aria del sistema $\mathfrak{R}$ y $x_1, \ldots, x_n$ son variables con recorrido en $\mathfrak{B}$, entonces $D\mathbf{r}(x_1, \ldots, x_n)$.
1. Si $D\varphi$, entonces también $D\neg\phi$.
2. Si $D\varphi$ y $D\psi$, entonces $D(\varphi \wedge \psi)$ y $D(\varphi \vee \psi)$.
3. Si $Df(x_1, \ldots, x_n)$ para todos los valores admisibles de las variables $x_1, \ldots, x_n$, entonces $D\forall x_1 \ldots \forall x_n f(x_1, \ldots, x_n)$ y también, por lo tanto, $D\exists x_1 \ldots \exists x_n f(x_1, \ldots, x_n)$.
4. Si $DF(f)$ para todo functor bien definido $f = f(x_1, \ldots, x_n)$, entonces $D\forall fF(f)$ y $D\exists fF(f)$.

Hoy llamamos 'functor $n$–ádico' a un símbolo que, acompañado de $n$ términos (esto es, expresiones gramaticalmente aptas para designar objetos), forma un nuevo término. Por ejemplo, 'log' es un functor monádico que, junto a un término que designe un número, forma un término que designa el logaritmo de ese número; '+' es un functor diádico que, acompañado de términos que designen dos números, forma un término que designa la suma de esos dos números. No me parece que en la cláusula 4 'functor' pueda entenderse en

[13] He aquí una traducción literal del texto de Zermelo: "Suponemos dado un dominio $B$ (o en general una pluralidad de dominios $B_1, B_2, \ldots$) así como una sistema $R$ de *relaciones fundamentales* de la forma $r(x, y, z, \ldots)$, donde las variables $x, y, z, \ldots$ pertenecen respectivamente a los dominios $B$" (1929, p. 342). En el simbolismo matemático actual, las expresiones '$B_1, B_2, \ldots$' y '$x, y, z, \ldots$' indicarían que Zermelo está hablando de unas secuencias infinitas de dominios y de variables. En tal caso, las generalizaciones universales y existenciales a que se refiere la cláusula II.3 de la definición de 'definit' irían precedidas por infinitos cuantificadores. Pero no creo que Zermelo haya contemplado aquí tan inusitada formación sintáctica. Tampoco creo que haya pretendido establecer una correspondencia biunívoca entre su lista de variables y su lista de dominios, como sugiere la frase "las variables $x, y, z, \ldots$ pertenecen respectivamente a los dominios $B$". Por otra parte, no habría inconveniente en admitir una partición del dominio $\mathfrak{B}$ en infinitos subdominios, y sólo en aras de la brevedad he omitido mencionar esta posibilidad en mi paráfrasis.

esta acepción. Significa más bien lo que hoy llamaríamos un *predicado*, esto es, un símbolo que, acompañado de un número idóneo de términos, forma una proposición. Así, lo que esa cláusula nos está diciendo es que se puede formar una proposición bien definida cuantificando sobre variables predicativas. Ello implica que la teoría de conjuntos no se deja formalizar en un cálculo predicativo de primer orden.[14]

La misma revista —*Fundamenta Mathematicae*— en que apareció el artículo de Zermelo (1929) publicó poco después una nota polémica de Thoralf Skolem (1930). Skolem observa que la definición de 'definit' propuesta por Zermelo, en la medida en que es aceptable, equivale a la suya, publicada mucho antes (Skolem 1922), y se sorprende de que Zermelo no la mencione. Conforme a la definición de Skolem, una aseveración bien definida (*eine definite Aussage*) es cualquier expresión finita construida a partir de aseveraciones elementales de la forma $a \in b$, o $a = b$, mediante una o más de las cinco operaciones siguientes: conjunción, disyunción, negación, cuantificación universal —"validez en cada caso"— y cuantificación existencial —"validez al menos en un caso" (Skolem, SWL, p. 139).[15] En otras palabras, una aseveración bien definida es una aseveración bien formada de un cálculo predicativo de primer orden con identidad y un primitivo único, el predicado diádico '$\in$'. Al permitir la cuantificación sobre funciones proposicionales, la cláusula 4 de la definición de Zermelo ciertamente asigna a 'definit' una extensión más amplia que la definición de Skolem. La crítica de Skolem se dirige principalmente contra esta cláusula. Según él, no está claro el significado de la expresión '$F(f)$' que allí figura. Designa, al parecer, una función de funciones proposicionales; pero Zermelo no explica como debe entenderse este concepto. Skolem reclama una mayor precisión sobre esto. En efecto, según él, si el concepto de función proposicional se piensa con tal amplitud que abarque el concepto de función de funciones proposicionales, puede generarse la paradoja de Russell. Por otra parte, si sólo se admiten funciones de funciones proposicionales construidas a partir de éstas mediante las cinco operaciones lógicas arriba mencionadas, los

[14] Para los lectores que no estén familiarizados con la lógica moderna, doy una versión de cálculo predicativo de primer orden en el Apéndice IX.

[15] Nótese que esta definición concuerda con la de Weyl (1910), que Skolem no menciona.

axiomas de Zermelo (1908a) combinados con la definición de 'definit' de Zermelo (1929) no autorizan la formación de ningún conjunto que no sea posible ya en virtud de esos mismos axiomas combinados con la definición más simple de 'definit' de Skolem (1922).[16]

### *1.8.3 El Axioma de Reemplazo*

El sistema de axiomas para la teoría de conjuntos conocido por la sigla ZF —con Z de Zermelo y F de Fraenkel— se distingue del sistema original de Zermelo (1908a) en cuanto excluye el Axioma de Selección (cuya inclusión genera el sistema llamado ZFC, con C de *choice*) y contiene un axioma nuevo, el Axioma de Reemplazo, propuesto —independientemente— por Fraenkel y por Skolem en 1922.[17]

Fraenkel (1922) justifica la necesidad del Axioma de Reemplazo así: Si $Z_0$ es el conjunto infinito cuya existencia postula el Axioma VII y (para cada $n > 0$) designamos con $Z_n$ al conjunto potencia $\mathcal{P}Z_{n-1,}$ cuya existencia resulta de la aplicación reiterada del Axioma IV a $Z_0$, es fácil ver que los axiomas de Zermelo no garantizan la existencia del conjunto infinito $\{Z_0, Z_1, \dots\}$.

16 Skolem, SWL, pp. 276–77. Skolem 1930 concluye con "una observación de orden general" que anticipa aspectos esenciales de la famosa argumentación de Gödel 1931. "Si para caracterizar las funciones proposicionales se establecen axiomas que expresen ciertos requisitos de reproducción como las cláusulas [1–4] de Zermelo, tales axiomas se dejan concebir [ . . . ] como 'aseveraciones aritméticas'" (Skolem, SWL, p. 278). Sea $O$ el más pequeño modelo concebible del dominio de funciones proposicionales caracterizado por tales axiomas. En virtud del teorema de Löwenheim-Skolem *(vide infra*, p. 98), $|O| = \aleph_0$. Supongamos que la teoría axiomática de conjuntos de Zermelo (mejorada con su definición de 'definit') tiene un modelo $B$. Entonces, $B$ contiene un conjunto infinito numerable $Z_0$ y existe una biyección de $Z_0$ en $O$. Sea $S_n$ la función proposicional de una variable que dicha biyección asigna a $n \in Z_0$. Sea $\Sigma(x)$ la función proposicional de una variable que es verdadera para $x = n \in Z_0$ si y sólo si $S_n(n)$ es falsa. $\Sigma(x)$ obviamente difiere de $S_n(x)$ para cada $n \in Z_0$ y por lo tanto no pertenece al modelo $O$. Por otra parte, $\Sigma(x)$ es una función proposicional bien definida conforme a la exigencia que Zermelo (1929, p. 341) impone a este concepto. "En otras palabras, parece imposible caracterizar mediante una dada axiomática finita un concepto de 'definit' que concuerde con [dicha exigencia]" (Skolem, SWL, p. 279).

17 Mirimanoff (1917, p. 49) postuló que si un conjunto $A$ es equinumeroso con una colección de conjuntos $B$, entonces $B$ también es un conjunto; este postulado equivale al Axioma de Reemplazo. Pero la obra interesantísima de Mirimanoff no fue reconocida

Por lo tanto, si el cardinal del continuo es menor que $\aleph_\omega$, no es posible establecer en el sistema de Zermelo que hay conjuntos cuyo cardinal sea igual o mayor que $\aleph_\omega$.[18] El Axioma VII, claro está, se puede reformular de modo que asegure la existencia del conjunto $\{Z_0, Z_1, \ldots\}$, pero el procedimiento utilizado para definirlo podría entonces aplicársele a él mismo, caracterizando así nuevos conjuntos cuya existencia, aunque manifiestamente plausible, no estaría garantizada por el sistema. Fraenkel propone, en cambio, una solución general de la dificultad que consiste en adoptar el axioma siguiente:

AXIOMA DE REEMPLAZO. Si *M* es un conjunto y cada elemento de *M* se reemplaza con [lo que Zermelo llama] *una cosa del dominio* $\mathfrak{B}$, *M* se convierte en otro conjunto.

Este enunciado que habla de reemplazos y conversiones tiene más de receta que de axioma, pero poco más tarde Fraenkel (1925, p. 254) ofrecerá una formulación más rigurosa: Si $\varphi$ es una función —en el sentido de Fraenkel (1922a) explicado en la Sección 1.8.2— y *M* es un conjunto, entonces también es un conjunto la imagen de *M* por $\varphi$, esto es, $\{\varphi(x): x \in M\}$. Ocurre, empero, que —como demostrará von Neumann (1928a, p. 377)— esta aseveración se deduce de los axiomas I–V de Zermelo (1908a) y por lo tanto no basta para establecer que existe un conjunto cuyo cardinal es igual o mayor que $\aleph_\omega$. Pero esta debilidad se remedia recurriendo a un concepto de función menos estrecho. Una función $\varphi$ en el sentido de Fraenkel tiene que estar definida en cierto conjunto *M*. En cambio, según von Neumann (1928a, p. 376), la función $\varphi$ mencionada en el Axioma de Reemplazo debe estar libre de esta restricción y admitir como argumento cualquier objeto susceptible de serlo que pertenezca al dominio de la teoría.

Así fortalecido, el Axioma de Reemplazo concuerda precisamente con el propuesto por Skolem al Congreso de Matemáticos Escandinavos de 1922, en la misma ponencia en que introdujo la definición de *definit* que cité en la Sección 1.8.2 (Skolem 1922):

---

por los estudiosos de la teoría de conjuntos antes de 1925, cuando fue citada por von Neumann (*vide infra*, nota 20).

18 Fraenkel hizo esta objeción a Zermelo en una carta fechada el 6 de mayo de 1921 (citada en Moore 1982, p. 263 n. 10).

> Sea $U$ una aseveración bien definida (*definite*) que vale para ciertos pares ordenados $\langle a,b \rangle$ de objetos pertenecientes al dominio $\mathfrak{B}$ de modo que para cada $a$ hay a lo sumo un $b$ tal que $U$ es verdad. Según $a$ recorre los elementos de un conjunto $M_a$, $b$ recorre todos los elementos de un conjunto $M_b$.
>
> (Skolem, SWL, pp. 145s.)

Skolem exhibe la insuficiencia del sistema original de Zermelo mediante el mismo ejemplo aducido por Fraenkel, a saber, el conjunto $\{Z_0, Z_1, \ldots\}$ arriba descrito; pero Skolem da una demostración rigurosa de la insuficiencia de ese sistema para garantizar la existencia de dicho conjunto. Para ello introduce el importante concepto de *rango* (*Stufe*) de un conjunto, que caracteriza como sigue. Sean $\mathcal{P}^n M$ y $\bigcup^n M$ los conjuntos formados reiterando $n$ veces la operación de formar, respectivamente, el conjunto potencia y el conjunto unión de un conjunto dado $M$ (Axiomas IV y V de Zermelo). Sea $\mathcal{P}^0 M = \bigcup^0 M = M$. Diremos que el conjunto $M$ es *de primer rango* si hay un entero $n \geq 0$ tal que $\bigcup^n M = \varnothing$. (Recordando que $Z_0 = \{\varnothing,\{\varnothing\},\{\{\varnothing\}\}, \ldots\}$, advertimos que $\bigcup^n Z_0 = Z_0$ para todo $n \geq 0$; por lo tanto, $Z_0$ no es un conjunto de primer rango). Para cada entero $k > 1$, diremos que el conjunto $M$ es *de k-ésimo rango* si $M$ no es de $(k–1)$-ésimo rango pero hay un entero $n \geq 0$ tal que todos los elementos de $\bigcup^n M$ son conjuntos de $(k–1)$-ésimo rango. (Nótese que, conforme a esta estipulación, $Z_0$ es un conjunto de segundo rango). Sea $B$ un dominio en que satisface los siete axiomas de Zermelo. En virtud del Axioma VII, $B$ contiene un conjunto infinito que llamaremos, como siempre, $Z_0$. Supongamos que $B$ contiene además el conjunto $\{Z_0, Z_1, \ldots\}$. Sea $B'$ el subdominio formado por todos los conjuntos de primer o segundo rango que hay en $B$. Se puede verificar fácilmente que $B'$ también satisface los axiomas de Zermelo. Obviamente, $Z_0$ pertenece a $B'$, puesto que es un conjunto de segundo rango, pero $\{Z_0, Z_1, \ldots\}$ es un conjunto de tercer rango y por lo tanto no puede pertenecer a $B'$. Como los axiomas de Zermelo se cumplen en $B'$, es claro que no bastan para garantizar la existencia del conjunto $\{Z_0, Z_1, \ldots\}$. En una nota, Skolem esboza una prueba de que el Axioma de Reemplazo basta para remover la deficiencia (Skolem, SWL, p. 146n.).

### *1.8.4 Aportes de von Neumann*

John von Neumann fue uno de los matemáticos más versátiles y fecundos del siglo XX. Su teoría matemática de los juegos y sus ideas sobre el diseño de computadoras han revolucionado el mundo de los negocios. Su aporte a la sistematización matemática y la clarificación conceptual de la mecánica cuántica, aunque muy discutida, también ha tenido una influencia enorme.[19] Contribuyó decisivamente a la teoría de conjuntos con una serie de trabajos publicados cuando tenía entre 20 y 25 años. En esta sección presentaré su innovadora concepción de los ordinales (1922/23) y las ideas centrales de su sistema axiomático (1925, 1928), aunque sin dar una presentación completa del mismo. En el Apéndice VIII explico cómo empleó su teoría de los ordinales para justificar la definición por inducción transfinita.

Como sabemos, Cantor llegó a definir los ordinales como "tipos de orden". Dos conjuntos ordenados tienen el mismo "tipo de orden" si son "similares", esto es, si hay entre ellos un isomorfismo de órdenes. Un ordinal es el tipo de orden de un conjunto bien ordenado. Von Neumann considera que este procedimiento es "algo vago" y propone reemplazarlo por otro, "basado en operaciones unívocas con conjuntos". La definición de Cantor se deja traducir, por cierto, a términos estrictamente conjuntistas: "El ordinal de un conjunto bien ordenado es la clase de todos los conjuntos similares a él". Pero si esto es un ordinal, no hay ninguna garantía de que existan conjuntos de ordinales y el intento de formarlos puede generar paradojas.

19 Sobre la teoría de los juegos, véase von Neumann 1928b, von Neumann y Morgenstern 1944. Sobre la mecánica cuántica, véase en particular von Neumann 1927a, 1932; Birkhoff y von Neumann 1936. La principal contribución de von Neumann al diseño de la computadora moderna consistió nada menos que en la idea misma de *software*, esto es, la idea de que las instrucciones para resolver un problema deben registrarse temporalmente en la memoria electrónica del aparato —al igual que los datos del problema, aunque premunidas de un codificación que diferencie estas dos clases de *input*— en vez de incorporarlas en la configuración fija de sus circuitos eléctricos. La idea es simple, pero genial, y aparentemente no se le había ocurrido a ninguno de los matemáticos e ingenieros que trabajaron en el diseño de computadoras electrónicas antes de que von Neumann entrase en este campo en 1944. En las primeras computadoras, construidas para las fuerzas armadas de Gran Bretaña y los Estados Unidos durante la Segunda Guerra Mundial, había que reconfigurar los circuitos cada vez que se quería abordar otro tipo de problemas. Cf. Pollack 1982, p. 19.

La definición de ordinal de von Neumann es una aplicación brillante del Axioma de Reemplazo que Fraenkel acababa de publicar (von Neumann lo cita). Sea $\langle M,<\rangle$ un conjunto ordenado. Si $x \in M$, el conjunto $\{y \in M: y < x\}$ —abreviado $A(x,M)$— es el segmento de $x$ en $\langle M,<\rangle$. Supongamos que $\langle M,<\rangle$ es un conjunto bien ordenado. Una *enumeración* (*Zählung*) de $M$ es una aplicación $f$ que asigna a cada $x \in M$ el conjunto $f(x) = \{f(y): y \in A(x,M)\}$. En otras palabras, el valor de $f$ para cada $x \in M$ es el conjunto de todos los valores de $f$ correspondientes a los elementos que preceden a $x$ en $\langle M,<\rangle$. Obsérvese que $f(x)$ es un conjunto en virtud del Axioma de Reemplazo, puesto que $A(x,M)$ lo es en virtud del Axioma de Separación. También $\{f(x): x \in M\}$ es un conjunto (por el Axioma de Reemplazo), que llamaré provisionalmente 'el ordinal de $\langle M,<\rangle$ determinado por la enumeración $f$'.

Cualquiera que sea el conjunto bien ordenado $\langle M,<\rangle$, si $x_0$ es su primer elemento, $f(x_0) = \varnothing$. Por lo tanto, si $x_1$, $x_2$ y $x_3$ son respectivamente, el segundo, tercer y cuarto elemento de $\langle M,<\rangle$, tenemos que

$$\begin{aligned} f(x_1) &= \{\varnothing\} \\ f(x_2) &= \{\varnothing,\{\varnothing\}\} \\ f(x_3) &= \{\varnothing,\{\varnothing\},\{\varnothing,\{\varnothing\}\}\}. \end{aligned}$$ [20]

Visiblemente, el conjunto bien ordenado $\langle M,<\rangle$, admite una y sólo una enumeración $f$, la cual determina un ordinal único, que llamaremos en adelante *el ordinal de* $\langle M,<\rangle$. Si $\langle M,<\rangle$ es un conjunto bien ordenado y $f$ es su enumeración, entonces $f(x) \notin f(x)$ para todo $x \in M$. Si $x, y \in M$ y $x < y$, entonces $f(x) \subset f(y)$. Von Neumann (1922/23) ofrece demostraciones detalladas de estas proposiciones y luego establece la siguiente importante caracterización de los ordinales:

> $P$ es un ordinal si y sólo si (i) $P$ es un conjunto de conjuntos, ordenable por inclusión; (ii) su orden por inclusión es un buen orden; (iii) si $\xi \in P$, $\xi = A(\xi, P)$ (cada elemento de $P$ es su propio segmento en $\langle P,\subset\rangle$).

[20] Este ejemplo lo daba ya Mirimanoff (1917, p. 46), quien descubrió antes que von Neumann esta forma de construir los ordinales. Von Neumann conocía la obra del matemático ruso y la menciona expresamente (1925, p. 230n.).

Usando esta caracterización von Neumann prueba que si $P$ es un ordinal, $P$ es el conjunto de todos los ordinales diferentes de $P$ e incluidos en $P$; que, si $P$ y $Q$ son dos ordinales diferentes uno de los dos está incluido en el otro, y que si $U$ es un conjunto cualquiera de ordinales, la relación de inclusión $\subset$ determina un buen orden en $U$. Una segunda caracterización, derivable de la anterior, es ésta: $P$ es un ordinal si y sólo si todo elemento de $P$ es un ordinal incluido en $P$. Con estos recursos, von Neumann puede demostrar fácilmente que los ordinales definidos a su manera satisfacen estos teoremas clásicos de la teoría cantoriana del buen orden: Dos conjuntos bien ordenados son similares si y sólo si sus ordinales respectivos son idénticos; un conjunto bien ordenado es similar a un segmento de otro conjunto bien ordenado si y sólo si el ordinal del primero está incluido en el ordinal del segundo; si $\langle M,<\rangle$ es un conjunto bien ordenado hay un y sólo un ordinal que, ordenado por inclusión, es similar a $\langle M,<\rangle$, a saber, el ordinal de $\langle M,<\rangle$.

Gracias a que concibe los ordinales como conjuntos de cierta índole von Neumann puede identificar los cardinales infinitos con ordinales, como se hace con los finitos. Digamos que el ordinal $\alpha$ es un *cardinal* si todo otro ordinal equinumeroso con $\alpha$ es mayor que $\alpha$ (von Neumann 1928, p. 731). En otras palabras, un ordinal es un cardinal si es el primer ordinal de su numerosidad. Como cada ordinal finito es más numeroso que su predecesor, todos son cardinales. En cambio, entre los ordinales infinitos sólo es un cardinal el primero de cada clase cantoriana. Los cardinales infinitos como los define von Neumann corresponden, pues, precisamente a los alephs de Cantor. Ahora bien, si un conjunto $M$ admite un buen orden $<$, $M$ es equinumeroso con el ordinal de $\langle M,<\rangle$; por lo tanto, bajo el Axioma de Selección cada numerosidad está representada por un aleph, o sea, un cardinal en el sentido de von Neumann. Pero si no se acepta este axioma, será conveniente disponer de una segunda definición de 'cardinal' restringida a los conjuntos que no pueden ser bien ordenados (cf. Levy 1979, p. 83).

"La construcción axiomática de la teoría general de los conjuntos" es el tema de la tesis doctoral sometida por von Neumann en 1925 a la Universidad de Budapest (en húngaro). Sabemos que en agosto de 1923 había enviado una versión alemana a Zermelo.[21] Fraenkel recordará treinta años después

[21] Véase la carta de von Neumann a Zermelo del 15 de agosto de 1923 en Meschkowski 1967, pp. 271–73.

que “alrededor de 1922–23” la redacción de *Mathematische Zeitschrift* lo consultó sobre “un largo manuscrito de un autor para él desconocido, Johann von Neumann, titulado ‘Die Axiomatisierung der Mengenlehre’, el cual era su futura disertación doctoral, que la revista publicaría recién en 1928.” Fraenkel dice que, aun sin entenderlo del todo, pudo reconocer *ex ungue leonem* y dio una opinión favorable. A la vez, invitó a von Neumann a una entrevista en la cual le aconsejó que “preparara el terreno para la comprensión de un ensayo tan técnico con otro más informal que destacase el nuevo enfoque del problema y sus principales consecuencias.”[22] Siguiendo este consejo, von Neumann redactó “Eine Axiomatisierung der Mengenlehre” (1925). Este artículo, que he tomado de base para la exposición siguiente, no deriva teoremas, pero en cambio explica la motivación de los axiomas y contiene importantes consideraciones de interés filosófico que faltan en la presentación formal (1928).

Von Neumann empieza describiendo dos modos de reaccionar ante las paradojas de la teoría de conjuntos. El primero, representado por Russell, Brouwer, Weyl y el húngaro König, “se propone la meta de establecer la ciencia exacta entera sobre una nueva base evidente para todos, desde la cual pueda nuevamente alcanzarse lo ‘correcto’ en matemáticas y teoría de conjuntos, pero excluyendo de antemano lo contradictorio gracias a la fundamentación intuitiva inmediata” (1925, p. 220). La impresión que dejan estos autores es “aniquiladora”: mientras Russell hace descansar toda la matemática en el cuestionable “Axioma de Reducibilidad”, Weyl y Brouwer rechazan consecuentemente la mayor parte de ella como carente de sentido y critican a fondo los métodos elementales de inferencia empleados normalmente en sus demostraciones. La segunda reacción, iniciada por Zermelo, retiene los métodos lógicos tradicionales pero descarta el concepto ingenuo de *conjunto*. Para reemplazarlo, se postulan axiomas en que figura la palabra ‘conjunto’ pero sin atribuirle un significado: “Por ‘conjunto’ se entiende aquí [ . . . ] sólo una cosa de la que no se sabe ni se quiere saber más que lo que al respecto se sigue de los postulados. Éstos tienen que formularse de

[22] Las dos citas de Fraenkel están tomadas de una carta suya a Ulam, transcrita en parte en Moore 1982, p. 264.

tal modo que de ellos se sigan todos los teoremas deseados de la teoría de conjuntos cantoriana, pero no las antinomias" (p. 220).

Von Neumann dice que su escrito pertenece a la segunda corriente. Con todo, el término matemático tradicional que caracteriza mediante axiomas no es 'conjunto', sino 'función', utilizado luego para *definir* 'conjunto'. Obviamente cualquier conjunto $M$ (en el sentido intuitivo del término) puede representarse de un modo inequívoco mediante su función característica, que asigna un valor fijo (por ejemplo, 1) a todos los objetos que pertenecen a $M$ y otro valor fijo (por ejemplo, 0) a todos los objetos que no pertenecen a $M$. El sistema de von Neumann se apoya en esta observación. Sus axiomas se refieren a dos clases de objetos, llamados informalmente 'funciones' y 'argumentos' (el título formal es 'cosas-I' y 'cosas-II' —*I. Dinge* y *II. Dinge*— pero lo evitaré en mis explicaciones). Hay dos argumentos distinguidos, $A$ y $B$ (que hacen el papel de 0 y 1 en el ejemplo anterior). Si $a$ y $b$ son argumentos, $\langle a,b\rangle$ es un argumento, el par ordenado cuyo primer elemento es $a$ y cuyo segundo elemento es $b$. (Repitiendo la operación de formar pares, se obtiene el $n$-tuplo $\langle a_1, \ldots, a_n\rangle = \langle\langle a_1, \ldots, a_{n-1}\rangle, a_n\rangle$). Si $f$ es una función y $a$ es un argumento, $[f,a]$ es un argumento, el valor de $f$ en $a$. Puede haber argumentos que a la vez son funciones —los llamaré 'funciones-argumentos' (el título formal es *I.II. Dinge*)— y es esencial que haya funciones que no pueden ser argumentos, esto es, objetos cuyo nombre sólo puede ocupar el primer lugar —nunca el segundo— en una expresión de la forma $[f,a]$.

Un *dominio* (*Bereich*) es una función $f$ tal que, para todo argumento $x$, o bien $[f,x] = A$, o bien $[f,x] = B$. Si el dominio $f$ es una función-argumento, von Neumann dice que $f$ es un *conjunto*. Introduce la abreviatura $a \in f$ para decir que $f$ es una función y $a$ es un argumento tal que $[f,a] \neq A$. Obsérvese que para usar esta abreviatura no se requiere que $[f,a] = B$; en otras palabras, es lícito escribir $a \in f$ aunque $f$ no sea un dominio. Por otra parte, si $f$ y $g$ son funciones, la expresión '$f \in g$' tiene sentido si y sólo si $f$ es una función-argumento. Evidentemente, si $f$ es un dominio pero no es un conjunto —esto es, si $f$ no es una función-argumento—, no puede existir un dominio $g$ tal que $f \in g$. En tal caso, diré que $f$ es un dominio propio.[23]

[23] Entiéndase 'propio' como lo contrario de 'impropio'. Más castizo sería decir 'dominio propiamente tal' pero la expresión es incómodamente larga. En la terminología de Gödel (1940), comúnmente aceptada en la literatura matemática actual, un objeto $k$ es

Intuitivamente, un dominio propio $f$ se puede ver como una colección de elementos —los argumentos $x$ tales que $[f,x] = B$— que no es a su vez elemento de ninguna colección. Gracias a que admite dominios que no son conjuntos von Neumann elude las paradojas clásicas sin coartar la libertad del matemático para concebir objetos y darles nombre. Como veremos, el sistema contiene un criterio preciso para determinar qué dominios no son conjuntos. Intuitivamente se lo puede expresar así: el dominio $f$ es un dominio propio si y sólo si la totalidad de los argumentos $x$ tales que $x \in f$ no es menos numerosa que la totalidad absoluta de los argumentos.

Los axiomas de von Neumann forman cinco grupos. El grupo I de "axiomas introductorios" postula la existencia de los argumentos $A$ y $B$, fija la sintaxis de las expresiones $[x,y]$ y $\langle x,y \rangle$, y estipula que las funciones $f$ y $g$ son idénticas si $[f,x] = [g,x]$ para todo argumento $x$. Los grupos II y III —"axiomas de construcción aritmética y lógica"— sirven principalmente para asegurar que las todas las expresiones debidamente formadas con designadores (constantes o variables) de argumentos y funciones y los símbolos [, ⟨, ], ⟩, =, ∀, puedan reducirse a ciertas formas normales (como explicaré en las dos próximas notas).

El grupo II postula que existen (1) la función identidad, esto es, una función $i$ tal que, para todo argumento $x$, $[i,x] = x$; (2) una función constante para cada argumento $u$, esto es, una función $c_u$ tal que para cada argumento $x$, $[c_u,x] = u$; (3 y 4) las proyecciones sobre el primer y sobre el segundo elemento de cada par ordenado, esto es, funciones $p_1$ y $p_2$ tales que, para todo par $\langle x,y \rangle$, $[p_1,\langle x,y \rangle,] = x$ y $[p_2,\langle x,y \rangle,] = y$; (5) una función $h$ tal que, para cada par ordenado $\langle f,x \rangle$ cuyo primer elemento es una función-argumento, $[h,\langle f,x \rangle] = [f,x]$. Postula además que, para cualesquiera funciones dadas $f$ y $g$, existen (6) la función producto $f \times g$ tal que, para todo argumento $x$, $[f \times g,x] = \langle [f,x],[g,x] \rangle$, y (7) la función compuesta $f \circ g$ tal que, para todo argumento $x$, $[f \circ g,x] = \langle [f,[g,x]] \rangle$.[24]

---

una *clase* (*class*) si es el conjunto vacío, $k = \{x: x \neq x\}$, o cumple la condición $\exists x(x \in k)$. Si $\forall x(k \notin x)$, decimos que la clase $k$ es una *clase propia* (*proper class*).

24 Gracias a los axiomas del grupo II von Neumann (1928, pp. 670ss.) puede demostrar el siguiente Teorema de Reducibilidad: Sea $\tau$ una expresión formada según las reglas del grupo I con nombres de argumentos y funciones, los símbolos [, ], ⟨, ⟩, y $n$ variables $\xi_1, \dots ,\xi_n$. Sea $\tau(a_1, \dots ,a_n)$ la expresión que se obtiene al reemplazar consistentemente

El grupo III postula (1) que hay una función $f$ tal que, para todo par ordenado $\langle x,y\rangle$, $\langle x,y\rangle \in f$ si y sólo si $x = y$; (2) que, si $f$ es una función, hay una función $h$ tal que para todo argumento $x$, $x \in h$ si y sólo si para todo argumento $y$, $\langle x,y\rangle \notin f$; y (3) que, si $f$ es una función, hay una función $h$ tal que para todo argumento $x$, $[h,x] = y$ cuando quiera que $y$ es el único argumento tal que $\langle x,y\rangle \in f$.[25]

El grupo IV demanda más atención. Los dos axiomas de que consta sirven, respectivamente, para distinguir a las funciones-argumentos de los argumentos que no son funciones y de las funciones que no son argumentos. No obstante la simetría de sus propósitos, su alcance es muy diferente. El Axioma IV.1 es prescindible, pues no habría ningún inconveniente formal en suponer que todos los argumentos son a la vez funciones, esto es, que cualquier objeto de la teoría puede nombrarse en primer lugar en una expresión encerrada por corchetes (aunque esta suposición contraría evidentemente nuestras ideas intuitivas de *argumento* y *función*). En cambio, el Axioma IV.2 es la pieza central del sistema de von Neumann y de él se deducen el Axioma de Separación, el Axioma de Reemplazo y el Axioma de Selección. He aquí una traducción literal de estos axiomas (seguida de una paráfrasis entre paréntesis):

IV.1 Hay una cosa-II $a$ tal que una cosa-I $x$ es una cosa-I/II si y sólo si $[a,x] \neq A$.
(Hay una función $a$ tal que un argumento $x$ es una función-argumento si y sólo si $x \in a$.)

---

en $\tau$ la variable $\xi_k$ por un nombre del argumento $a_k$ $(1 \leq k \leq n)$. Entonces existe una función $f$ tal que, para todo $n$-tuplo de argumentos $\mathbf{x} = \langle x_1, \ldots, x_n\rangle$, si $\tau(\mathbf{x})$ designa un argumento, $[f,\mathbf{x}] = \tau(\mathbf{x})$.

25 El grupo III garantiza que, si $\Phi(\mathbf{a})$ expresa una condición impuesta a un $n$-tuplo de argumentos $\mathbf{a} = \langle a_1, \ldots, a_n\rangle$, habrá siempre una función $f$ tal que $\Phi(\mathbf{a})$ si y sólo si $[f,\mathbf{a}] \neq A$ (esto es, si y sólo si $\mathbf{a} \in f$). El Axioma III.1 lo asegura si $\Phi(\mathbf{a})$ envuelve identidades y III.2 si envuelve cuantificación universal o existencial, mientras que III.3 implica que cualquier objeto $y$ determinado implícitamente por una condición $[f,\langle x,y\rangle] \neq A$ tiene también una representación explícita $y = [h,x]$.

IV.2 Una cosa-II *a* no es una cosa-I/II si y sólo si hay una cosa-II *b* tal que para cada cosa-I *x* existe un *y* que cumple las condiciones $[a,y] \neq A$ y $[b,y] = x$.
(Una función *a* no es una función-argumento si y sólo si hay una función *b* tal que para todo argumento *x*, $[b,y] = x$ para algún $y \in a$.)

Como se puede ver, el Axioma IV.2, con toda su tremenda fuerza, es sumamente simple. Restringido a cosas-II (funciones) que sean dominios lo que dice es esto: Un dominio *a* es un dominio propio— y no un mero conjunto— si y sólo si hay una función *b* que lo aplica *sobre* el universo de todos los argumentos. Recordando el distingo tardío de Cantor entre pluralidades inconsistentes y consistentes (Capítulo 1.6), podemos decir que en el sistema de von Neumann una pluralidad es "demasiado grande" para ser recogida como elemento en otra pluralidad si, pero solamente si es "tan grande" como la pluralidad de todos los objetos ordinarios (conjuntos y elementos de conjuntos). Pero en este sistema un objeto así "tan grande" admite una interpretación natural como función que no es a su vez argumento de otra, y no hace falta entender que los objetos que *esa* función admite como argumentos no pueden "existir todos conjuntamente" (Cantor, GA, p. 443; citado en el Capitulo 1.6).

Los tres axiomas del grupo V —"axiomas del infinito"— equivalen precisamente al Axioma del Infinito, el Axioma de Unión y el Axioma del Conjunto Potencia. Sólo el primero difiere significativamente en su formulación del respectivo axioma de Zermelo, pues, aunque postula como éste la existencia de un conjunto enumerablemente infinito, lo construye de otro modo. Doy una paráfrasis de este axioma:

V.1 Existe una función-argumento *a* con las siguientes propiedades: Hay una función-argumento *x* tal que $x \in a$. Si *y* es una función-argumento tal que $y \in a$, hay una función-argumento $z \neq y$ tal que $z \in a$ y para todo *x*, $x \in y$ implica $x \in z$.

Las funciones-argumentos de que aquí se habla no tienen que ser conjuntos, pero si todas lo fuesen el conjunto *a* cuya existencia se postula admitiría la caracterización siguiente: (i) *a* es un conjunto no vacío; (ii) algunos elementos de *a* son conjuntos; (iii) todo conjunto que es un elemento de *a* es parte de

otro conjunto diferente que también es un elemento de *a*.

Von Neumann hace hincapié en que el grupo V sólo es indispensable para la teoría de los cardinales infinitos, pero que los primeros cuatro grupos de axiomas bastan para fundar la teoría de los conjuntos finitos y los números naturales (enteros no negativos), así como una parte significativa de la teoría del continuo, inclusive la definición de los números reales como secuencias de Cauchy de racionales, los teoremas de convergencia de secuencias y series, la teoría de las funciones continuas, el álgebra, el análisis y la integral de Riemann. Pero sin el grupo V "carecen de sentido el Teorema de Weierstraß sobre el supremo (de conjuntos numéricos, no secuencias), puesto que no puede haber conjuntos de cosas-II; asimismo el concepto general de función, el buen orden del continuo, la integral de Lebesgue" (von Neumann 1925, p. 237).

Después de presentar los axiomas y bosquejar rápidamente el orden en que la teoría de conjuntos se deduce de ellos (en su 1928), von Neumann desarrolla algunas consideraciones escépticas sobre su significado y utilidad (1925, pp. 229–40). Ellas giran esencialmente en torno al hecho de que la teoría ofrecida no es categórica. En la Sección 1.8.1 indiqué que una teoría axiomática se dice *categórica* si todos sus modelos son isomórficos. Como un axioma nada dice sobre la naturaleza individual de los objetos a que se refiere y sólo puede especificar la *estructura* resultante de sus relaciones mutuas, una teoría categórica caracteriza su tema con la máxima precisión alcanzable por un sistema axiomático. Von Neumann muestra que la teoría determinada por sus axiomas no es categórica y sostiene que ninguna teoría axiomática de conjuntos puede serlo. Según él, esto resulta de un teorema demostrado por Löwenheim (1915) y generalizado por Skolem (1920, 1922). Conforme a ese teorema, si una teoría $T$ formulable en un cálculo predicativo de primer orden tiene un modelo infinito, $T$ tiene un modelo $M_\kappa$ con cardinal $\kappa$ para cada $\kappa \geqslant \aleph_0$. Obviamente, $M_\kappa$ no puede ser isomórfico a $M_\lambda$ si $\kappa \neq \lambda$. El Teorema de Löwenheim-Skolem se aprecia mejor a la luz de la semántica de Tarski y me referiré a él en el Capitulo 3.2. Por eso, de las ideas presentadas por von Neumann en la última parte de su artículo de 1925 mencionaré aquí una sola, que no envuelve dicho teorema y que nos ayudará a entender el trabajo de Zermelo (1931) que estudiaremos en la Sección 1.8.5.

Los *Fundamentos de la geometría* de Hilbert incluyen desde la segunda edición un "axioma de completud",[26] destinado a asegurar la categoricidad de la teoría. En virtud del mismo, un modelo de la teoría no puede estar incluido como parte propia en otro modelo que preserve las relaciones geométricas entre los elementos del primero.[27] Inspirándose en el ejemplo de Hilbert, Fraenkel propuso completar la teoría axiomática de conjuntos con el siguiente Axioma de Limitación:

$AL_F$ Aparte de los conjuntos que forzosamente tienen que existir en virtud de los axiomas, no existen otros conjuntos.

Von Neumann (1925, p. 230) observa que esta formulación es vaga y la reemplaza por otra más precisa (ajustada, además, a su terminología). Sea $\Sigma$ un modelo de la teoría de von Neumann, esto es, un sistema de cosas-I y cosas-II ("argumentos" y "funciones"), en el cual se distinguen dos cosas-I, *A* y *B*, y las operaciones $\langle\ ,\ \rangle$ y $[\ ,\ ]$ están definidas de manera que se cumplan todos los axiomas de los grupos I–V. $\Sigma'$ es un subsistema de $\Sigma$

[26] Pido excusas por este neologismo feísimo, pero firmemente arraigado entre los especialistas de habla castellana (algunos prefieren decir 'completitud', que no suena mejor). Corresponde al sustantivo alemán 'Vollständigkeit' y al inglés 'completeness', que designan justamente la calidad de lo es o está *completo* (*vollständig*, *complete*). Más eufónico hubiera sido quizás 'complenitud', ligado al verbo latino *compleo* (de donde *completus*, 'completo') del mismo modo que 'plenitud' está ligado con el verbo *pleo* ('llenar'). Pero los profesores de lógica ya no estudian lenguas clásicas.

[27] "V.2 (*Axioma de completud lineal*). El sistema de los puntos de una recta con sus relaciones de orden y congruencia no admite una ampliación que preserve las relaciones entre los elementos anteriores así como las propiedades fundamentales que se derivan de los Axiomas I–III de orden lineal y de congruencia y V.1" (Hilbert, GG, p. 30; el Axioma V.1 es el Postulado de Arquímedes). Sin este axioma la teoría de Hilbert no es categórica, pues el modelo numérico mencionado en la Sección 1.8.1 incluye una parte propia que satisface los axiomas restantes si se mantiene la interpretación de los primitivos y la definición pitagórica de distancia, a saber, el conjunto de los tríos ordenados de números algebraicos. Esta estructura no es isomórfica a la primera, pues el conjunto de los números algebraicos es numerable, mientras que $\mathbb{R}$ no es numerable. Obsérvese que en virtud del Axioma V.2, el sistema de los puntos construibles con regla y compás, estudiado por Euclides, *no es un modelo* de la teoría de Hilbert, 2ª edición, a pesar de que satisface los Axiomas I–IV y V.1; porque dicho sistema es una parte propia del espacio homeomorfo a $\mathbb{R}^3$ estudiado por Descartes y los geómetras modernos, que es un modelo de esta teoría.

si $\Sigma'$ consta de un sistema I′ de cosas-I y un sistema II′ de cosas-II, todas las cuales pertenecen también a $\Sigma$, y $A'$ y $B'$ son cosas distinguidas de I′ (posiblemente diversas de $A$ y $B$), y las operaciones $\langle\ ,\ \rangle'$ y $[\ ,\ ]'$ definidas en $\Sigma'$ son las restricciones de $\langle\ ,\ \rangle$ y $[\ ,\ ]$ a I′ y II′ (es decir, $\langle f,x\rangle' = \langle f,x\rangle$ y $[x,y]' = [x,y]$ si $f$ pertenece a II′ y $x$ e $y$ pertenecen a I′). Diré que el subsistema $\Sigma'$ es un *submodelo* si los axiomas de los grupos I–V también se cumplen si se entiende que se refieren a las cosas-I de I′, las cosas-II de II′, las operaciones $\langle\ ,\ \rangle'$ y $[\ ,\ ]'$ y los objetos distinguidos $A'$ y $B'$.[28] Usando este vocabulario el Axioma de Limitación puede reformularse así:

$AL_N$ Si $\Sigma$ es un modelo de los axiomas de los grupos I–V y $\Sigma' \neq \Sigma$ es un subsistema de $\Sigma$, $\Sigma'$ no es un submodelo.

Von Neumann señala que la relación ‘ser un subsistema de’ que aquí figura no puede asimilarse a la relación ‘$\subseteq$’ definible en su teoría ($u \subseteq v$ si y sólo si $u$ y $v$ son funciones y para todo argumento $x$, $x \in u$ implica que $x \in v$), por cuanto el modelo $\Sigma$ de que se habla necesariamente contiene funciones (cosas-II) que no son argumentos (cosas-I), y por lo tanto no puede ser él mismo una función con la que otra función tenga la relación ‘$\subseteq$’. Parecería pues que hay que entender la relación ‘ser un subsistema de’ en su acepción conjuntista ingenua, que tendríamos que dar por supuesta. El intento de rescatar la teoría de Cantor mediante la caracterización axiomática de sus conceptos fundamentales se vería entonces abocado al fracaso. Como solución de esta dificultad, von Neumann adelanta la idea que me interesa presentar. Consideremos un modelo $P$ de la teoría de von Neumann que comprende un sistema $I_P$ de argumentos y un sistema $II_P$ de funciones, con operaciones $\langle\ ,\ \rangle_P$ y $[\ ,\ ]_P$ y argumentos distinguidos $A_P$ y $B_P$. Supongamos que todas las cosas-I y las cosas-II del anterior modelo $\Sigma$ están contenidas en $I_P$ y que para toda cosa-II $f$ de $\Sigma$ hay una función $\phi$ en $II_P$ tal que, para toda cosa-I $x$ de $\Sigma$, $[\phi,\langle f,x\rangle_P]_P$ es idéntico al valor $[f,x]$ determinado por la

[28] Menciono dos objetos distinguidos $A'$ y $B'$ (o, respectivamente, $A$ y $B$) por seguir a von Neumann, pero en rigor para obtener un modelo de sus axiomas basta distinguir una sola cosa del tipo I. En efecto, los axiomas de von Neumann mencionan un solo argumento distinguido ($A$); el segundo ($B$) aparece en las definiciones de ‘dominio’ y ‘conjunto’, términos que no se utilizan en los axiomas.

operación [ , ] en $\Sigma$. Entonces $\Sigma$ es un dominio en $P$ y los subsistemas de $\Sigma$ tienen con $\Sigma$ la relación '$\subseteq$' (en $P$).

> Así habríamos sobreimpuesto a $\Sigma$ una "teoría de conjuntos superior" $P$, en la que también son argumentos cosas que en $\Sigma$ no pueden serlo. Esto no es absurdo de suyo. Si aceptamos como argumentos en un sistema nuevo $P$ a los conjuntos que eran "demasiado grandes" para ser argumentos [en $\Sigma$], podemos siempre esquivar las antinomias si admitimos a su vez los conjuntos "aún más grandes" (esto es, demasiado grandes en $P$) constituidos por aquéllos, pero los declaramos inadmisibles como argumentos.
> (Von Neumann 1925, p. 231)

Aunque la nueva idea enseguida se revela impotente para fundar una teoría categórica,[29] abre a la reflexión matemática perspectivas insondables. Veremos ahora cómo Zermelo las despliega —con decidido espíritu cantoriano (y leibniziano)— en la presentación de su segundo sistema de axiomas para la teoría de conjuntos.

29 Sea **N** la teoría determinada por los cinco grupos de axiomas de von Neumann y **NL** la determinada por **N** y el axioma $AL_N$. **NL** es categórica por definición, pero habría que ver si es consistente (relativamente a **N**). Para ello hay que formular condiciones necesarias y suficientes para que un subsistema de un modelo de **N** sea un submodelo y hallar un modelo de **N** ninguno de cuyos subsistemas las satisfaga. Von Neumann considera imposible formular tales condiciones sin incurrir en un círculo vicioso. Por ejemplo, el Axioma II.1 es verdadero en un subsistema $\Sigma'$ de un modelo $\Sigma$ de **N** si y sólo si hay en $\Sigma'$ una función $f$ tal que, *para todo argumento $x$ en $\Sigma'$*, $[f,x] = x$. Según von Neumann no es lícito cuantificar sobre los argumentos del subsistema $\Sigma'$ en el enunciado de las condiciones que deben servir para caracterizarlo. Podemos, claro está, cuantificar sobre los argumentos (y funciones) de $\Sigma$, pero entonces obtenemos condiciones *suficientes*, mas no *necesarias* para que un subsistema de $\Sigma$ sea un submodelo. Von Neumann da una lista de estas condiciones y muestra que todo modelo $\Sigma$ de **N** contiene un submodelo *mínimo* $\Sigma_\mu$ que las cumple todas. Entonces ningún subsistema de $\Sigma_\mu$ cumple todas esas condiciones, pero cualquiera de ellos podría ser un submodelo aunque no las cumpla.

### *1.8.5 Zermelo (1930)*

En el artículo "Sobre números-límite y dominios de conjuntos" (1930), Zermelo presenta un nuevo modo de concebir el tema mismo de la teoría de conjuntos. Formula un sistema de axiomas que se inspira en el suyo de 1908 aunque difiere de él significativamente. El sistema no es categórico y Zermelo, que ve en ello una virtud (véase la nota 6), procede a establecer verdades generales sobre sus posibles modelos. Cada modelo $M$ está caracterizado por dos parámetros: el cardinal de su "base" de *Urelemente* (véase la nota 5) y lo que Zermelo llama su "característica", esto es, el primer ordinal mayor que todos los ordinales realizados en $M$.[30] Zermelo demuestra que dos modelos $M$ y $M'$ son isomórficos si y sólo si tienen bases equinumerosas y la misma característica, y que si la característica de $M$ es mayor que la de $M'$, $M$ contiene un submodelo isomórfico a $M'$. Las características inducen, pues, un buen orden entre las clases de equivalencia (por isomorfismo) de modelos con bases equinumerosas. Por otra parte, si dos modelos tienen la misma característica, uno de ellos será en todo caso isomórfico a una parte del otro, aunque sus bases no sean equinumerosas. La investigación "metateórica" de los modelos de la teoría axiomática se conduce en el estilo y con los conceptos del conjuntismo "ingenuo" de Cantor. Pero si existe un modelo $M_\alpha$ con característica $\alpha$, la metateoría restringida a modelos de característica menor que $\alpha$ se deja establecer (en $M_\alpha$) como una consecuencia de la teoría axiomática. (Por otra parte, como el ordinal $\alpha$ *no está realizado* en $M_\alpha$, de todos modos hay que rebasar los recursos disponibles en este modelo para que dicha metateoría restringida pueda siquiera ser descrita como tal.)

El sistema axiomático propuesto por Zermelo en 1930 reproduce sus axiomas de 1908 con dos notables excepciones: 1º Se excluye el Axioma de Selección (VI), porque "tiene otro carácter que los demás y no sirve para

[30] Zermelo (1930) concibe los ordinales como tipos de orden, a la manera de Cantor. Así entendido, el ordinal $\alpha$ está realizado en el modelo $M$ si existe en $M$ un conjunto bien ordenado con tipo de orden $\alpha$. Pero si concebimos a los ordinales a la manera de von Neumann la expresión '$\alpha$ está realizado en $M$' puede entenderse en sentido literal: 'el conjunto $\alpha$ existe en $M$'. Evidentemente, si $\alpha$ está realizado en $M$ en esta segunda acepción también lo está en la primera. Por otra parte, si hay en $M$ un conjunto bien ordenado $a$ con tipo de orden $\alpha$, los axiomas de Zermelo (1930) aseguran que el ordinal de von Neumann $\alpha$ también existe en $M$ ($\alpha$ es el alcance de la enumeración de $a$).

delimitar los dominios [de los modelos]” y Zermelo ahora lo considera como “un principio lógico universal presupuesto por toda nuestra investigación” (1930, p. 31). 2º Se excluyen los asertos existenciales absolutos, es decir, la postulación del conjunto vacío ∅ (I, primera parte) y el Axioma del Infinito (VII).[31] Se incluyen, pues, aseveraciones equivalentes a los Axiomas de Determinación, Separación, Formación de Pares (I, última parte), Conjunto Potencia y Unión. Se incluye además el Axioma de Reemplazo y un nuevo “Axioma de Fundación” (mejor conocido en la literatura actual como “Axioma de Regularidad”), en virtud del cual no puede haber ninguna secuencia infinita $a_1, a_2, \ldots$ tal que $a_{n+1} \in a_n$ para todo $n \geq 1$ (por ende, no hay cadenas infinitas de la forma $\ldots \in a_3 \in a_2 \in a_1$, ni ciclos de la forma $a \in b \in \ldots \in a$).[32] Zermelo designa cada uno de estos axiomas por letras que —con una excepción— corresponden a la inicial del respectivo nombre alemán. Llama *ZF* al sistema BAPUVE de los primeros seis, y *ZF′* al sistema completo BAPUVEF (F por *Fundierung*, ‘fundación’).[33] La presentación de los axiomas termina con esta importante observación metodológica:

> Llamamos *“dominio normal”* a un dominio de “conjuntos” y “elementos primordiales” (“*Urelemente*”) que satisfaga nuestro sistema *ZF′* con respecto a la “relación fundamental” $a \in b$. A los “dominios” de este género, sus “elementos”, “subdominios”, “uniones” e “intersecciones” los trataremos conforme a los conceptos y axiomas universales de la teoría

31 El Axioma del Infinito reaparecerá al final, fortalecido a ultranza, como postulado “metateórico”. Con respecto a ∅ tenemos sólo la indicación, en el Axioma del Conjunto Potencia, de que cualquiera que sea el conjunto $m$, el respectivo conjunto $\mathcal{P}m$ no está vacío, sino contiene en cada caso el mismo *Urelement* arbitrariamente escogido $u_0$ (“que hace las veces del ‘conjunto vacío’”). Esta no es una aseveración absoluta de existencia, pero implica que, si existe un conjunto, existe al menos un objeto que no es un conjunto.

32 Gracias a que da por supuesto el Axioma de Selección, Zermelo puede enunciar elegantemente su Axioma de Fundación así: Todo conjunto no vacío $x$ contiene un conjunto $y$ tal que para todo $z \in x$, $z \notin y$. Cf. Mendelson 1958.

33 Cito el Axioma de Separación A (*Aussonderung*): “Toda función proposicional $f(x)$ separa en cada conjunto $m$ un subconjunto $m_f$ que contiene todos los elementos $x$ tales que $f(x)$ es verdad” (Zermelo 1930, p. 30). En una nota al pie de esa página Zermelo explica que la función proposicional $f(x)$ —así como la función de reemplazo en el Axioma de Reemplazo— puede ser “enteramente *arbitraria*”, de modo que “todas las consecuencias derivadas de limitar estas funciones a una clase especial caducan desde el punto de vista aquí adoptado ”.

> de conjuntos, exactamente como [si fuesen] *conjuntos*, ya que no se distinguen materialmente (*sachlich)* de éstos en ningún aspecto esencial; pero los llamaremos "dominios" y no "conjuntos" para diferenciarlos de los "conjuntos" que son elementos del dominio bajo consideración.
>
> (Zermelo 1930, p. 31)

En la terminología de Zermelo, un conjunto bien ordenado $m$ es una *secuencia fundamental* (*Grundfolge*) si el primer elemento de $m$ es un *Urelement* y todo otro elemento $m$ es igual al conjunto de los elementos precedentes. Así, una secuencia fundamental es —como un ordinal de von Neumann— un conjunto bien ordenado por la relación $\in$. Por lo demás, si convenimos con Zermelo en que $\varnothing$ esté representado en cada dominio normal por un cierto *Urelement* fijo $u_0$ (véase nota 31), los ordinales de von Neumann son secuencias fundamentales. Zermelo demuestra que si un dominio normal $D$ contiene un *Urelement* $u$ y un conjunto bien ordenado $r$ con tipo de orden $\rho$, $D$ contiene también una secuencia fundamental $g_\rho$ similar a $r$ cuyo primer elemento es $u$.[34] La totalidad de las secuencias fundamentales con idéntico primer elemento $u$ contenidas en un dominio fundamental $D$ constituyen un subdominio bien determinado $G_u$ de $D$ y sus respectivos tipos de orden forman un segmento bien determinado $Z_\delta$ de la serie de los ordinales, con tipo de orden $\delta$. Pero $G_u$ no es uno de los conjuntos de $D$, ni hay en $D$ un conjunto bien ordenado con tipo de orden $\delta$, sino que $\delta$ es el primer ordinal mayor que todos los ordinales realizados en $D$. "De otro modo, se obtendría la conocida Antinomia de Burali-Forti" (Zermelo 1930, p. 33). El ordinal $\delta$, así descrito, es la *característica* o *número-límite* del dominio normal $D$. A continuación explicaré dos propiedades esenciales de tales números-límite establecidas por Zermelo. Para aligerar la exposición, adoptaré las definiciones de ordinal y cardinal de von Neumann (Sección 1.8.4); de acuerdo con ellas, "el segmento $Z_\delta$ de la serie de los ordinales determinado por el ordinal $\delta$" es precisamente el ordinal $\delta$ y un ordinal es un cardinal si no es equinumeroso con un ordinal menor.

[34] Por el Axioma de Reemplazo: sea $u$ el reemplazante del primer elemento de $r$; réemplacese cualquier otro elemento de $r$ por el conjunto de los reemplazantes de sus predecesores.

(I) Si $\delta$ es la característica de un dominio normal, $\delta$ es un cardinal regular. Esto quiere decir que si $\alpha$ es un ordinal menor que $\delta$, y $\sigma_\alpha$ es un subconjunto de $\delta$ tal que $\langle\sigma_\alpha,<\rangle$ es similar a $\langle\alpha,<\rangle$, entonces $|\alpha| \prec |\delta|$ y hay un ordinal $\beta < \delta$ tal que $\sigma_\alpha \subseteq \beta$.[35]

(II) Se puede definir una aplicación $f$ de la serie de los ordinales en sí misma, tal que (*a*) para cualesquiera ordinales $\xi$ y $\zeta$, $f(\xi) < f(\zeta)$ si y sólo si $\xi < \zeta$; (*b*) si $\alpha$ es un ordinal límite, $\lim_{\xi<\alpha} f(\xi) = f(\lim_{\xi<\alpha}\xi)$ y (*c*) si $\delta$ es la característica de un dominio normal, $f(\delta) = \delta$. Esta aplicación queda definida por las tres condiciones siguientes: (i) $f(0) = 0$; (ii) $f(\xi + 1) = |\mathcal{P}f(\xi)|$ para cada ordinal sucesor $\xi + 1$, y (iii) $f(\alpha) = \lim_{\xi<\alpha} f(\xi)$ para cada ordinal límite $\alpha$.[36]

35 Demostración de (I). Sea $\delta$ la característica de un dominio normal $D$. Si $\delta$ no es un cardinal hay un ordinal $\gamma < \delta$, tal que $|\gamma| = |\delta|$. En tal caso, $\gamma$ está realizado en $D$. Por el Teorema de Cantor, $|\mathcal{P}\gamma| \succ |\gamma| = |\delta|$. $\mathcal{P}\gamma$ existe en $D$ (por el Axioma del Conjunto Potencia) y admite un buen orden similar a algún ordinal $\beta$. Entonces $|\beta| = |\mathcal{P}\gamma| \succ |\delta|$, de suerte que $\delta$ no es mayor que todos los ordinales realizados en $D$, contra la hipótesis. Por lo tanto, $\delta$ es un cardinal. Supongamos ahora que $\sigma$ es un subconjunto de $\delta$ tal que $\langle\sigma,<\rangle$ es similar a $\langle\alpha,<\rangle$ para algún ordinal $\alpha < \delta$ y que hay un $\xi \in \sigma$ tal que para todo $\beta < \delta$, $\beta \leq \xi$. En tal caso, $\bigcup\sigma$ existe en $D$ (por el Axioma de Unión). Pero $\bigcup\sigma$, por ser la unión de un conjunto de ordinales, también es un ordinal, a saber, $\lim_{\xi\in\sigma}\xi = \delta$. Por lo tanto, $\bigcup\sigma$ no puede existir en $D$. En consecuencia, no existe un $\sigma \subset \delta$ que cumpla la condición señalada. Por lo tanto, $\delta$ es un cardinal regular.

Con un argumento similar al primero de los anteriores, Zermelo prueba también que si $\delta$ es la característica de un dominio normal $D$, $\delta = \aleph_\alpha$ para algún ordinal límite $\alpha$. En efecto, por ser un cardinal, $\delta = \aleph_\alpha$ para algún ordinal $\alpha$. Pero si $\alpha$ fuera un ordinal sucesor, $\alpha = \xi + 1$, el cardinal $\aleph_\xi$ estaría realizado en $D$ y también, por lo tanto, el cardinal $|\mathcal{P}\aleph_\xi| \succcurlyeq \aleph_{\xi+1} = \aleph_\alpha = \delta$, y $\delta$ no podría ser la característica de $D$. Por consiguiente, $\alpha$ es un ordinal límite.

36 Demostración de (II). Sea $\delta$ la característica de un dominio normal $D$. Probaré por inducción transfinita (Apéndice VIII) que si $\alpha$ es un ordinal menor que $\delta$, $f(\alpha) < \delta$. Supongamos que $f(\xi) < \delta$ para todo $\xi < \alpha$. Entonces, para todo $\xi < \alpha$, $\mathcal{P}f(\xi)$ existe en $D$, de modo que $f(\xi + 1) = |\mathcal{P}f(\xi)| < \delta$. Por lo tanto, si $\alpha$ es un ordinal sucesor, $f(\alpha) < \delta$. Si $\alpha$ es un ordinal límite, sea $\alpha' = \{f(\xi): \xi < \alpha\}$. Como $\alpha$ está realizado en $D$, $\alpha'$ existe en $D$ por el Axioma de Reemplazo y $\bigcup\alpha'$ existe en $D$ por el Axioma de Unión. Pero $\bigcup\alpha' = \lim_{\xi<\alpha} f(\xi) = f(\alpha)$, de modo que $f(\alpha) < \delta$, puesto que es un ordinal que está realizado en $D$. En la demostración de (I) se vio que $\delta$ no puede ser un ordinal sucesor. Por lo tanto es un ordinal límite y $f(\delta) = \lim_{\xi<\delta} f(\xi)$. Si $\delta < f(\delta)$ habría un $\alpha < \delta$ tal que $f(\alpha) > \delta$, contra lo que se acaba de probar. Por lo tanto, $\delta = f(\delta)$.

Las propiedades (I) y (II) caracterizan a los *cardinales inaccesibles*.[37] Se los llama así porque, si $\kappa$ es un cardinal de esta especie, no es posible "alcanzarlo" a partir de cardinales menores que $\kappa$ por las operaciones ordinarias de formar pares, uniones y conjuntos potencia. No es difícil entender por qué la característica $\delta$ de un dominio normal $D$ tenía que ser un cardinal inaccesible. $D$ es un modelo de la teoría $ZF'$, que postula la existencia de todo par, unión o conjunto potencia de conjuntos dados, y $\delta$ es el primer ordinal mayor que todos los realizados en $D$. Por lo tanto, si $\kappa$ y $\lambda$ son ordinales menores que $\delta$, $|\kappa\cup\lambda|$, $|\bigcup\kappa|$ y $|\mathcal{P}\kappa|$ están realizados en $D$ y por ende tienen que ser menores que $|\delta| = \delta$. Zermelo recurre reiteradamente a este argumento en la demostración de (I) y (II) (notas 35 y 36).

La parte central del trabajo de Zermelo (1930) investiga patrones estructurales discernibles en todo modelo de la teoría $ZF'$. Empieza con esta importante caracterización de los submodelos incluidos en un dado modelo:

LEMA: Sea $D'$ un subdominio del dominio normal $D$. $D'$ es un dominio normal si y sólo si los siguientes objetos pertenecientes a $D$ también pertenecen a $D'$: (i) cada elemento contenido en un conjunto perteneciente a $D'$ y (ii) cada conjunto que sólo contenga elementos pertenecientes a $D'$. Si $D'$ comprende toda la base de *Urelemente* de $D$, entonces $D' = D$.[38]

[37] Hausdorff (1914, p. 131) llamó "exorbitantes" a los ordinales que tienen la propiedad (I); hoy suele llamárselos "débilmente inaccesibles". Zermelo (1930, p. 35) observa acertadamente que bajo la Hipótesis Generalizada del Continuo, los números que tienen la propiedad (I) también tienen la propiedad (II): los cardinales débilmente inaccesibles también son inaccesibles (en el sentido "fuerte"). El primer ordinal infinito $\omega$ tiene, por cierto, las propiedades (I) y (II); algunos autores lo excluyen —por convención terminológica— de la clase de los cardinales inaccesibles.

[38] DEMOSTRACIÓN DEL LEMA. Hay que probar que los axiomas de la teoría $ZF'$ se cumplen en $D'$. Si $a$ y $b$ pertenecen a $D'$, $\{a,b\}$ es un conjunto perteneciente a $D'$ (por la condición (ii)). Si $m$ es un conjunto perteneciente a $D'$ y $m_\phi$ es el subconjunto de $m$ determinado por la condición $\phi$, $m_\phi$ pertenece a $D'$, pues $m_\phi$ es un conjunto de $D$ que sólo contiene elementos de $m$, los cuales pertenecen todos a $D'$ (por (i)). Si $m$ es un conjunto de conjuntos perteneciente a $D'$, todos sus elementos y los elementos de sus elementos pertenecen a $D'$ (por (i)); por ende, $\bigcup m$ pertenece a $D'$ (por (ii)). Como todos los elementos contenidos en un conjunto $m$ perteneciente a $D'$ también pertenecen a $D'$ (por (i)), todos los subconjuntos de $m$ existentes en $D$ existen en $D'$ (por (ii)); por lo tanto, $\mathcal{P}m$ pertenece a $D'$ (por (ii)). Análogamente, se puede mostrar que $D'$ satisface el Axioma de Reemplazo, si éste se refiere, como es natural, sólo a sustituciones de elementos de un conjunto existente en $D'$ por objetos pertenecientes a $D'$.

Premunido de este resultado, Zermelo establece que todo dominio normal $D$ con base $U$ y característica $\delta$ se descompone en una serie bien ordenada de estratos no vacíos y mutuamente disjuntos $Q_\alpha$ ($\alpha < \delta$), tales que un objeto $x$ del dominio $D$ pertenece a $Q_\alpha$ siempre que (i) para cada $\xi < \alpha$, $x \notin Q_\xi$ y (ii) para cada $y \in x$, hay un $\xi < \alpha$ tal que $y$ pertenece a $Q_\xi$. La descomposición del dominio normal $D$ en dichos estratos se apreciará mejor si prestamos atención a los "segmentos" $D_\alpha = \bigcup_{\beta<\alpha} Q_\beta$ ($0 < \alpha < \delta$). Tales segmentos pueden definirse así:

($S_1$) $D_1 = U = Q_0$.

($S_\alpha$) Si $0 < \alpha < \delta$, y $\alpha$ es un ordinal límite, $D_\alpha = \bigcup_{\xi<\alpha} D_\xi$; si $\alpha$ no es un ordinal límite, $\alpha = \beta+1$ y $D_{\beta+1}$ contiene todos los objetos de $D$ cuyos elementos pertenezcan a $D_\beta$.

Si estipulamos que $Q_0 = D_1 = U$ y $Q_\alpha = D_{\alpha+1} \setminus D_\alpha$ ($0 < \alpha < \delta$), es claro que los estratos $Q_\alpha$ ($\alpha < \delta$) determinan una partición del dominio $D_\delta = \bigcup_{\xi<\delta} D_\xi$. Se comprueba fácilmente que $D_\delta$ cumple las condiciones del lema. Como la base $U$ está incluida en $D_\delta$, tenemos que $D_\delta = D$.

Especial interés presenta la estratificación del dominio normal $D$ cuando es un "dominio unitario" (*Einheitsbereich*), esto es, cuando su base $U$ consta de un solo *Urelement*. Sea, pues, $|U| = 1$ y $f$ la aplicación definida en (II). Entonces, para cada $\alpha < \delta$, tenemos que (i) $|D_\alpha| = |f(\alpha)|$; (ii) si $x$ es un conjunto perteneciente a $D_\alpha$, $|x| \prec |D_\alpha|$; (iii) en el estrato correspondiente $Q_\alpha$ hay algún conjunto $x$ tal que $|x| = |D_\alpha|$; (iv) cada subdominio de $D_\alpha$ es un conjunto en $D_{\alpha+1}$, y (v) si $\alpha$ es un ordinal límite, todos los segmentos $D_\xi$ ($\xi < \alpha$) y sus respectivos subdominios son conjuntos en $D_\alpha$. Además, $|D| = \delta$ y cualquier subdominio de $D$ que sea menos numeroso que $D$ es un conjunto perteneciente a $D$. Según esto, si $D$ es un dominio normal unitario, $D$ satisface el Axioma IV.2 de von Neumann, según el cual una colección de objetos es "demasiado grande" para ser un conjunto sólo si es tan numerosa como todo el universo de objetos. Zermelo subraya que un modelo de $ZF'$ que tenga más de un *Urelement* no satisface ese axioma de von Neumann (1930, p. 38). Muestra asímismo que cualquier dominio normal $D$ admite una "estratificación canónica", cuyos "segmentos canónicos" se definen por las estipulaciones siguientes:

($C_1$) $D_1 = U = Q_0$ (como $S_1$).

($C_\alpha$) Si $0 < \alpha < \delta$, y $\alpha$ es un ordinal límite, $D_\alpha = \bigcup_{\xi<\alpha} D_\xi$; si $\alpha$ no es un ordinal límite, $\alpha = \beta+1$ y $D_{\beta+1}$ contiene cada objeto $x$ de $D$ cuyos elementos pertenecen a $D_\alpha$ y cuyo cardinal $|x| \leqslant |f(\alpha)|$ —donde $f$ es la aplicación definida en (II).

Si $D$ es un dominio unitario, ($C_\alpha$) equivale a ($S_\alpha$), de modo que $D$ no admite otra estratificación que ésta. La estratificación canónica le sirve a Zermelo para establecer que las propiedades (I) y (II) son no sólo *necesarias* sino también *suficientes* para que un dominio sea "normal" (esto es, para que sea un modelo de *ZF′*). También se utiliza en la demostración de los teoremas de isomorfismo a que aludí al comienzo y enunciaré con mayor precisión ahora. Sea $U_k$ la base y $\delta_k$ la característica de un dominio normal $D_k$ ($k$ = 1, 2). *Teorema* I. Si $|U_1| = |U_2|$ y $\delta_1 = \delta_2$, $D_1$ y $D_2$ son isomórficos y una biyección $\varphi$: $U_1 \to U_2$ determina un isomorfismo único $D_1 \to D_2$. *Teorema* II. Si $|U_1| = |U_2|$ y $\delta_1 < \delta_2$, $D_1$ es isomórfico a un segmento canónico de $D_2$. *Teorema* III. Si $\delta_1 = \delta_2$, o bien $D_1$ es isomórfico a un subdominio (propio o impropio) de $D_2$ o bien $D_2$ es isomórfico a un subdominio (propio o impropio) de $D_1$. De la segunda parte del teorema I, Zermelo deduce fácilmente que el grupo de automorfismos de un dominio normal es isomórfico (como grupo) al grupo de permutaciones de su base.

En la última sección de su trabajo, Zermelo aborda "cuestiones de existencia, consistencia y categoricidad", bajo el supuesto de que el sistema *ZF′* no implica contradicciones. Este supuesto no se ha demostrado, pero puede aceptarse como plausible mientras no se derive una contradicción de los axiomas BAPUVEF (y ninguna ha sido derivada hasta la fecha). Zermelo da, entonces, por descontado que existe un dominio normal $D$ con característica $\delta$ y una base no especificada $U$. El teorema III de isomorfismo implica que si $U'$ es una parte propia de $U$ existe un dominio normal $D'$ con base $U'$. En cambio, como no todo ordinal es un "cardinal inaccessible" con las propiedades (I) y (II), una variación de la característica $\delta$ no genera automáticamente un nuevo dominio normal. Ni siquiera es claro que haya tales "cardinales inaccesibles", ni, si los hay, cuántos son.

El primer ordinal que reúne las propiedades (I) y (II) es $\omega$. Digamos que un objeto $m$ existente en $D$ *envuelve el infinito* si $m$ es un conjunto infinito

o si hay en $D$ $n$ conjuntos $a_1$, $a_2$ . . . , $a_n$ tales que $a_1$ es infinito y $a_1 \in a_2 \in \ldots \in a_n = m$ ($n \geq 1$). El segmento $D_\omega$ de la estratificación canónica de $D$ comprende precisamente todos los objetos de $D$ que no envuelven el infinito. $D_\omega$ es un dominio normal con característica ω. Por lo tanto, si $ZF'$ tiene un modelo cualquiera $D$ también tiene un modelo $D'$ con característica ω. La existencia de este modelo "finitista" —que Zermelo piensa que no será disputada ni siquiera por los peores adversarios de la teoría de conjuntos— basta al menos para probar que la teoría $ZF'$ es consistente. Pero, "justamente porque no contiene ningún conjunto infinito, no se lo puede ser aducir como un 'modelo' genuino de la teoría de conjuntos cantoriana" (1930, p. 44). El Axioma del Infinito de Zermelo (1908a), al postular la existencia de un conjunto similar a ω, supera los límites del modelo descrito. Si hay un modelo de $ZF'$ que satisfaga este axioma, esto es, un dominio normal $D$ en que el ordinal ω esté realizado, su característica es el primer cardinal regular κ que sea un punto fijo de la función $f$ definida en (II), en otras palabras, que satisfaga la ecuación $f(\kappa) = \kappa$. Pero es sensato preguntar: ¿existe acaso un ordinal $\kappa > \omega$ que tenga estas propiedades?

He aquí las consideraciones en que Zermelo basa su respuesta afirmativa a esta pregunta. Si $D$ es un dominio normal con característica κ, no existe en $D$ ningún conjunto bien ordenado similar a κ, ni en general ningún conjunto equinumeroso con κ. Por otra parte, la totalidad de las secuencias fundamentales con idéntico *Urelement* inicial existentes en $D$ tiene precisamente el cardinal $|\kappa|$ y forma una secuencia bien ordenada similar a κ. Pero ¿tenemos derecho a profesar que dicha totalidad *existe como tal*, si existe el dominio normal $D$? Zermelo recuerda que si se fija la característica y el cardinal de la base de un dominio normal $D$, la teoría $ZF'$ determina a $D$ "categóricamente" (es decir, $ZF'$ determina sin ambigüedades la clase de todos los dominios normales isomórficos a $D$). Por lo tanto, si adoptamos "la hipótesis general de que *todo dominio categóricamente determinado puede ser también concebido de algún modo como 'conjunto'*, es decir, que puede presentarse como elemento de un dominio normal (apropiado), resulta que a cada dominio normal corresponde otro superior con la misma base, a cada dominio unitario un dominio unitario superior, y así a cada 'número-límite'

π corresponde también un número-límite mayor π′" (1930, p. 46).[39] Esta reflexión se aplica también a cualquier totalidad categóricamente determinada de "números-límite", de modo que la serie de éstos números es tan ilimitada como la de los ordinales y a cada ordinal transfinito corresponde biunívocamente un cardinal inaccesible.

> Naturalmente, esto *no* es una consecuencia "demostrable" de los axiomas *ZF′*, ya que la situación aseverada trasciende cada dominio normal particular. Más bien hay que postular la *existencia de una serie ilimitada de números-límite* como un nuevo *axioma* para la "metateoría de conjuntos" (*für die "Meta-Mengenlehre"*).
>
> (Zermelo 1930, p. 46)

La nueva visión de la teoría de conjuntos arroja luz también sobre las paradojas en que se ceban los "antimatemáticos":

> Estas aparentes "contradicciones" descansan meramente en una confusión de la *teoría de conjuntos misma*, no determinada categóricamente por sus axiomas, con los *modelos* particulares que la representan: lo que en un modelo aparece como un "no-conjunto o superconjunto ultrafinito" (*"ultrafinite Un- oder Übermenge"*) es ya en el modelo superior siguiente un "conjunto" plenamente válido con número cardinal y tipo de orden, y constituye el cimiento sobre el cual se edifica el nuevo dominio. A la serie

[39] El lector habrá advertido que Zermelo formula su "hipótesis general" en dos versiones tales que la primera es plausible, pero la consecuencia buscada sólo se infiere de la segunda, a saber:

(1) Todo dominio categóricamente determinado puede ser también concebido *de algún modo* como "conjunto" ("*irgendwie* auch als 'Menge' aufgefaßt werden kann"—cursiva mía).

(2) Todo dominio categóricamente determinado puede presentarse como elemento de un dominio normal idóneo ("als Element eines (geeignet gewählten) Normalbereiches auftreten kann").

Aunque en el texto de Zermelo van unidos por la expresión "es decir" (*das heißt*), los asertos (1) y (2) no dicen lo mismo, y no es posible establecer su equivalencia lógica sin premisas adicionales. 'Conjunto' se usa en (1) en su significado "ingenuo" ordinario, y no está dicho que sólo un objeto que se llame 'conjunto' en un modelo de la teoría *ZF′* sea admisible como ejemplo ilustrativo de ese significado.

ilimitada de los ordinales cantorianos corresponde pues una doble serie igualmente ilimitada de modelos esencialmente distintos de la teoría de conjuntos, en cada uno de los cuales se expresa íntegra la teoría clásica. Las dos tendencias polarmente opuestas del espíritu que piensa, la idea del *progreso* creador y la de la *conclusión* abarcadora, que también subyacen a las "antinomias" kantianas, están representadas y reconciliadas simbólicamente en el concepto de la serie transfinita de los números, basada en el buen orden, la cual en su progreso sin límites no tiene una verdadera conclusión, pero sí estaciones, a saber, justamente esos "números-límite" que separan los tipos de modelo superiores de los inferiores. Y así las "antinomias" de la teoría de conjuntos, bien entendidas, conducen a la ciencia matemática no a un estrechamiento y mutilación, sino más bien a un despliege y enriquecimiento que aún es imposible abarcar con la mirada.

(Zermelo 1930, p. 47)

# 2

# CÁLCULOS

## 2.1 EL PROGRAMA DE HILBERT

Hacia 1920, David Hilbert, entonces en la cima de su fama, se embarcó resueltamente en un nuevo programa de fundamentación de las matemáticas. Más tarde, ante la Sociedad Filosófica de Hamburgo, explicará su propósito, lapidariamente, así:

> Con esta nueva fundamentación de las matemáticas persigo un fin importante: Al hacer de cada enunciado matemático una fórmula que pueda exhibirse en concreto y derivarse con rigor y al darle así a las conceptualizaciones e inferencias matemáticas una forma tal que resulten irrefutables y a la vez proporcionen una representación de la ciencia [matemática] entera, *yo quisiera eliminar definitivamente del mundo la cuestión de los fundamentos de las matemáticas*.
>
> (Hilbert 1931, p. 489; cursiva mía)

Su interés en el asunto no era nuevo. En 1899 había publicado la primera axiomatización adecuada de la geometría euclidiana.[1] Al año siguiente apareció su caracterización axiomática del cuerpo de los reales, que reemplaza —y vindica— las definiciones clásicas de los números reales como colecciones infinitas de racionales (Hilbert 1900).[2] Dos cuestiones de fundamentos encabezan la lista de problemas para el siglo XX que planteó ese mismo año al Primer Congreso Internacional de Matemáticos, reunido en París: el problema cantoriano de la cardinalidad del continuo y la consisten-

[1] En la Sección 1.8.1 me referí brevemente a esta obra. Mario Pieri (1899) publicó casi al mismo tiempo otra axiomatización de la geometría clásica, muy diferente de la de Hilbert, pero no menos idónea.

[2] Los axiomas de Hilbert definen el sistema de los reales como un cuerpo arquimédico completo (véase el Glosario, s.v. ‘cuerpo’). Las definiciones clásicas de Weierstraß, Méray, Dedekind y Cantor determinan sendos modelos (realizaciones) de esta especie de estructura. Como todos los modelos de un cuerpo arquimédico completo son isomórficos, dichas definiciones pueden aceptarse como equivalentes.

cia[3] de "los axiomas aritméticos", expresión con la que designa a los axiomas de la teoría de los reales. La importancia de este último problema se puede explicar así: Varias pruebas propuestas en el siglo XIX referían la consistencia de una teoría dudosa a la de otra teoría incuestionada. Lobachevsky, por ejemplo, había mostrado que las fórmulas trigonométricas de su geometría no-euclidiana podían generar una contradicción *sólo si* la trigonometría esférica *euclidiana* era inconsistente. El mismo Hilbert (1899) había probado que su axiomatización de la geometría euclidiana no era inconsistente a menos que la teoría de los números reales también lo fuera. Con ello, ancló las anteriores pruebas de consistencia *relativa* en la consistencia de la aritmética. Ésta tendría que demostrarse *directamente*, estableciendo la imposibilidad, no condicional, sino *absoluta* de inferir una contradicción de sus axiomas. Ante el Tercer Congreso Internacional, celebrado en Heidelberg, Hilbert (1904) explicó a grandes rasgos cómo habría que proceder a tal demostración directa. Esta ponencia, que manifiestamente subestima la dificultad de la empresa, adelanta ya algunos planteamientos y métodos característicos de su programa de los años 20.[4] Pero su renovado interés en el tema fue motivado al parecer por las críticas de Brouwer (1907, 1912) y de Weyl (1921) a la fundamentación conjuntista del análisis y el consiguiente rechazo por parte de estos eminentes matemáticos de la metodología matemática habitual.

La filosofía matemática de Brouwer y Weyl cae fuera de los límites de este libro. Pero para apreciar los motivos y la orientación del programa de Hilbert no es preciso conocer el pensamiento de estos autores, sino solamente la idea que se hacía al respecto el propio Hilbert (quien, al parecer, no

[3] Hilbert dice 'Widerspruchslosigkeit', literalmente, 'carencia de contradicción'. Un sistema de axiomas tiene esta propiedad si es imposible derivar de él aseveraciones contradictorias "mediante un número finito de inferencias lógicas" (Hilbert 1900a; GA, III, 300). En adelante, usaré siempre los sustantivos 'consistencia', 'inconsistencia' (así como los correspondientes adjetivos 'consistente', 'inconsistente') en esta acepción. Más adelante habrá oportunidad de precisar —de distintas maneras en diferentes casos— lo que se admite como "inferencia lógica". Otros autores llaman 'consistente' a una teoría, axiomatizada o no, si admite una realización o modelo. Para evitar confusiones, usaré siempre las palabras 'realizable', 'realizabilidad' para referirme a esta propiedad de las teorías (o de sus axiomas).

[4] Me referiré a Hilbert 1904 en la Sección 2.9.2, donde dispondremos de más elementos de juicio para apreciar sus propuestas.

puso demasiado empeño en entenderlos).[5] En su primer trabajo sobre el nuevo programa, se expresa así:

> [Weyl y Brouwer] buscan fundamentar la matemática arrojando por la borda todo lo que les parece incómodo y erigen una Dictadura de la Prohibición *à la* Kronecker. Con ello, empero, desmiembran y mutilan nuestra ciencia, y corremos el peligro de perder gran parte de nuestros más valiosos tesoros si seguimos a tales reformadores. Weyl y Brouwer proscriben los conceptos generales de número irracional y de función, inclusive de función aritmética, los números cantorianos de las clases superiores de números, etc.; también el teorema según el cual entre infinitos números naturales siempre hay uno menor que los demás, e incluso el principio lógico "Tertium non datur", por ejemplo, en la aseveración: o bien hay sólo un número finito de números primos o hay infinitos. Estos son ejemplos de proposiciones e inferencias prohibidas. Creo que así como Kronecker no logró suprimir el número irracional —Weyl y Brouwer permiten al menos conservar un muñón— tampoco Weyl y Brouwer tendrán éxito. No. Brouwer no es la revolución, como cree Weyl, sino sólo la repetición, con recursos viejos, de un amago de *putsch* que antes falló totalmente aunque fue intentado con mucha más decisión, y que, ahora que el poder del estado ha sido tan bien equipado y fortalecido por Frege, Dedekind y Cantor, está de antemano condenado al fracaso.
>
> (Hilbert 1922; GA, III, 159s.)

A mi modo de ver, el desacuerdo entre Kronecker y Cantor reside, en lo esencial, en esto: Kronecker exigía que toda definición matemática incluyera un método para decidir efectivamente a cuáles objetos se aplicaba y a cuáles no;[6] en cambio, Cantor no tenía inconveniente en reconocer que una colec-

5 El lector curioso encontrará en el Apéndice XIX un esbozo de las ideas de Brouwer contra las que reacciona Hilbert. Allí podrá comprobar que la parte *positiva* de la enseñanza de este autor no ha sido criticada por Hilbert, quien sólo buscó proteger a la matemática clásica contra su escepticismo, mediante una nueva e irrebatible fundamentación.

6 Kronecker difundía su concepción de las matemáticas desde su influyente cátedra en la Universidad de Berlín. De sus escritos, el más filosófico es tal vez el ensayo "Sobre el concepto de número" (1887), pero en él no figura la exigencia arriba mencionada. Pero el siguiente pasaje, tomado de un trabajo de Kronecker de 1882, la aplica muy claramente en un caso especial: "La citada definición de irreducibilidad carecerá de un fundamento seguro mientras no se indique un método mediante el cual se pueda decidir, ante una función determinada dada, si ella es o no irreducible conforme a dicha defi-

ción de objetos existe y está completamente determinada, aunque no haya un inventario de sus elementos o un método efectivo para generarlo. Al igual que Kronecker, Brouwer y sus partidarios rechazan este supuesto. Por eso, no pueden aceptar la idea moderna de aplicación (función) como correspondencia arbitraria entre conjuntos, ni la identificación de los números irracionales con secuencias de racionales (no convergentes en $\mathbb{Q}$), ni los ordinales transfinitos de Cantor. En ese rechazo se basa también la prohibición de invocar el Principio del Tercero Excluido o *Tertium non datur* en las demostraciones referentes a colecciones infinitas.[7] En efecto: supongamos que el mero enunciado de una condición no contradictoria $C$ no basta para determinar la colección de los objetos que son $C$; entonces, aunque se pruebe que es falso que todos los $C$ cumplen además una cierta condición $D$, no se habrá establecido que existe un objeto $x$ que es $C$ pero no es $D$; para certificar la existencia de tal objeto hay que presentarlo o, al menos, presentar un método para construirlo.

Hilbert, en cambio, entiende de muy distinta manera lo que un matemático puede y debe admitir como existente. Como le hace presente a Frege en la misma carta del 29 de diciembre de 1899 que cité en la Sección 1.8.1:

> Si los axiomas arbitrariamente estipulados, junto con todas sus consecuencias, no se contradicen entre sí, entonces son verdaderos y existen las cosas definidas por ellos. Ese es para mí el criterio de la existencia y de la verdad.
>
> (Frege, KS, p. 411)

En otras palabras, si el matemático caracteriza un sistema de objetos mediante una serie de condiciones mutuamente compatibles (esto es, mediante condiciones que no se contradigan explícitamente y no impliquen consecuen-

---

nición" (citado por O. Becker 1954, p. 328, de Kronecker, "Grundzüge einer arithmetischen Theorie der algebraischen Grössen", § 4).

7 El Principio del Tercero Excluido afirma que cualquier aseveración $p$ es verdadera o es falsa: no hay una tercera alternativa (*non datur tertium*); simbólicamente: $p \vee \neg p$. Aplicado a un predicado $\Gamma$ atribuible a los elementos de un conjunto caracterizado por un predicado $\Delta$, el Principio afirma que, o bien todos los elementos del conjunto $\{x\colon \Delta x\}$ son $\Gamma$, o bien alguno de ellos no es $\Gamma$; simbólicamente:

$$\forall x(\Delta x \rightarrow \Gamma x) \vee \exists x(\Delta x \wedge \neg\Gamma x)$$

cias contradictorias), tiene que reputar existente a cualquier objeto cuya pertenencia al sistema se infiera de dichas condiciones. Hilbert le da un giro un tanto distinto a esta idea en la citada ponencia ante el Congreso de París:

> Si a un concepto se le asignan características contradictorias, digo que *ese concepto matemáticamente no existe*. [ . . . ] Pero si se logra demostrar que las características asignadas al concepto no pueden conducir jamás a una contradicción mediante la aplicación de un número finito de inferencias lógicas, digo que con ello se ha demostrado la existencia matemática del concepto (por ejemplo, de un número o una función que cumple ciertos requisitos).
>
> (Hilbert 1900a; GA, III, 301–302)

Esta noción de existencia conceptual *ad usum mathematicorum* concuerda hasta cierto punto con lo que Cantor llamaba la “realidad inmanente” de los objetos matemáticos (Capítulo 1.6).[8] Como en este punto Hilbert discrepa absolutamente de Brouwer y los suyos, su polémica con ellos estaba de antemano destinada a ser un diálogo entre sordos.[9] Pero el propósito de Hilbert

8 Cantor exigía además que todo nuevo sistema introducido mediante una caracterización consistente estuviese firmemente arraigado en la tradición matemática. En sus palabras: los nuevos conceptos deben “estar en relaciones firmes, ordenadas mediante definiciones, con los conceptos previamente formados, ya existentes y probados” (Cantor, GA, p. 182; citado en la Sección 1.6). Podría decirse que Hilbert introduce una exigencia similar cuando compara un *problema* matemático nuevo —especialmente cuando lo suscita la investigación de la naturaleza— con “una ramita joven que sólo prospera y da fruto si se la injerta cuidadosamente según las reglas rigurosas del arte del jardinero en el tronco viejo, el seguro patrimonio de nuestro saber matemático” (1900a; GA, III, 293s.).

9 Van Dalen (1990) narra cómo Hilbert, que era el director de *Mathematische Annalen*, hizo sacar a Brouwer de la lista de colaboradores distinguidos impresa en la portadilla de la revista —donde figuró desde 1915 hasta 1928— porque, según le notificó, “no me es posible colaborar con usted, dada la incompatibilidad de nuestros puntos de vista sobre cuestiones fundamentales”. Una conferencia dictada por Hilbert en 1928 ante el Seminario Matemático de Hamburgo, contiene la siguiente indisimulada alusión a Brouwer: “Quitarle al matemático el *tertium non datur* sería como querer prohibirle el telescopio al astrónomo o el uso de los puños al boxeador. La prohibición de las aseveraciones existenciales y el *tertium non datur* viene a ser más o menos lo mismo que la renuncia a la ciencia matemática en general. [ . . . ] Me asombra que un matemático ponga en duda la validez rigurosa del modo de inferencia basado en el *tertium non datur*. Más me asombra que, al parecer, se haya formado ahora toda una comunidad de

no era quizás traerlos de vuelta al redil, sino sólo neutralizar de una vez por todas el efecto descorazonador que sus objeciones pudieran tener sobre los buenos matemáticos. En interés de la matemática misma, pero también de toda nuestra cultura, quiere establecer que “las proposiciones matemáticas de hecho son verdades inexpugnables y definitivas” (1922; GA, III, 162).[10] La incertidumbre que todavía prevalece al respecto se debe, según él, a que los estudios sobre los fundamentos de las matemáticas realizados hasta entonces

> no permiten reconocer un camino que haga posible formular cada cuestión relativa a los fundamentos de tal modo que tenga que ir seguida de una respuesta unívoca. Pero esto es justamente lo que reclamo: En principio, en asuntos matemáticos, no deben tener cabida las dudas, ni las medias verdades, ni las verdades esencialmente heterogéneas (*es soll in mathematischen Angelegenheiten prinzipiell keine Zweifel, es soll keine Halbwahrheiten und auch nicht Wahrheiten von prinzipiell verschiedener Art geben können*).
>
> (Hilbert 1922; GA, III, 157)

Como Hilbert (1899) había demostrado que su axiomatización de la geometría no es contradictoria a menos que el concepto clásico de número real lo sea, la certeza indisputable que Hilbert demanda se habrá logrado, a sus

---

matemáticos que lo hace. Pero sobre todo me asombra el hecho de que incluso entre matemáticos el poder de sugestión de un solo hombre ingenioso y temperamental (*die Suggestivkraft eines einzelnen temperamentvollen und geistreichen Mannes*) sea capaz de ejercer los efectos más excéntricos y más inverosímiles” (Hilbert 1928; GG[7], p. 307).

10 Más de una vez Hilbert ha descrito su programa como una cruzada cultural. Por ejemplo, en su alocución al Congreso de Matemáticos de Bologna:

> ¿En qué pararía la verdad de nuestro saber en general y la existencia y el progreso de la ciencia si ni siquiera en las matemáticas hubiese una verdad segura? Y en efecto, hoy por hoy, el escepticismo y el desánimo con respecto a la ciencia suelen expresarse incluso en la literatura especializada y en conferencias públicas. Esto es como una especie de ocultismo, que juzgo dañina. La Teoría de la Prueba hace imposible tal actitud y nos procura la convicción entusiasta de que al menos el entendimiento matemático no tiene límites y puede incluso rastrear las leyes del pensamiento mismo.
>
> (Hilbert 1928a, en GG[7], p. 323)

ojos, en cuanto dispongamos de una teoría consistente de los reales. Equiparando los reales con ciertos conjuntos infinitos de racionales, los conjuntistas intentaron fundar una teoría de los reales por lo menos tan consistente como la aritmética. Según Brouwer y Weyl, las paradojas de la teoría de conjuntos han puesto en evidencia la futilidad de ese intento. Hilbert no comparte esta manera de ver:

> El hecho de que el concepto de conjunto en el sentido más general no sea admisible sin más no excluye de ningún modo que el concepto de un conjunto de números enteros sea correcto. Y las paradojas de la teoría de conjuntos no pueden considerarse como una prueba de que el concepto de conjunto de números enteros lleva a contradicciones. Al contrario: todas nuestras experiencias matemáticas respaldan la corrección y la consistencia de este concepto.
>
> (Hilbert 1922; GA, III, 158)

Hilbert alude además a su propia axiomatización de la teoría de los reales (1900), que los caracteriza sin recurrir al concepto de conjunto de enteros.[11] Sus axiomas son suficientes para fundamentar el análisis. En este campo, "no obstante las más osadas y diversas combinaciones y el uso de los méto-

[11] Según Hilbert, gracias a su caracterización axiomática "las dudas que en general se han hecho valer contra la existencia del agregado (*Inbegriff*) de todos los números reales y de conjuntos (*Mengen*) infinitos en general pierden toda su justificación [ . . . ]: el conjunto de los números reales no tiene que concebirse ahora, digamos, como la totalidad (*Gesamtheit*) de las leyes que pueden gobernar la sucesión de los términos de una secuencia de Cauchy (*Fundamentalreihe*), sino más bien [ . . . ] como un sistema (*System*) de cosas cuyas relaciones mutuas están dadas por el sistema *finito y cerrado* de los Axiomas I–IV [que caracterizan un cuerpo arquimédico completo] y *sobre los cuales valen otras aseveraciones sólo si uno puede derivarlas de esos axiomas mediante un número finito de inferencias lógicas*" (1900; GG[7], p. 246). He puesto en cursiva la cláusula final (el original destaca sólo las tres palabras correspondientes a 'finito y cerrado') porque tendré que referirme a ella más adelante. En el pasaje paralelo de Hilbert 1900a está expresada en forma aún más drástica así: "El continuo [en la concepción axiomática es] un sistema de cosas cuyas relaciones mutuas están reguladas por los axiomas establecidos y con respecto a las cuales son verdaderos todos los hechos y solamente los hechos que se pueden deducir de los axiomas mediante un número finito de inferencias lógicas (*für welche alle und nur diejenige Tatsachen wahr sind, die durch eine endliche Anzahl logischer Schlüsse aus den Axiomen gefolgert werden können*)" (GA, III, 301).

dos más refinados, reina una total seguridad en las inferencias y una concordancia patente de todos los resultados" (Hilbert 1922; GA, III, 159), y por lo tanto se justifica aceptar los axiomas que cimientan esa seguridad y concordancia. "Disputar esa justificación sería privar de antemano a toda ciencia de la posibilidad de operar" (ibid.). Subsiste, con todo, el problema de probar la consistencia de los axiomas. Este es justamente el problema que Hilbert quiere resolver con su nuevo programa.

Contestando a la citada carta de Hilbert (del 29.XII.1899), Frege se pregunta qué medios hay para demostrar que ciertas propiedades o requisitos no se contradicen mutuamente. Responde así:

> El único que conozco es este: Exhibir un objeto que posea todas esas propiedades, indicar un caso en que se cumplan todos esos requisitos. No sería posible demostrar la consistencia (*Widerspruchslosigkeit*) por otra vía.
> (Frege, KS, p. 414)

Obviamente es inútil aducir la consistencia de un conjunto de axiomas como criterio de la existencia de un sistema de objetos caracterizado por él, si para cerciorarse de aquélla hay que tener a éstos por delante. Pero el programa de Hilbert se encamina a demostrar directamente la imposibilidad de inferir una contradicción de las teorías axiomáticas que le interesan. Ello presupone, por cierto, que se conozcan con precisión todas las formas de inferencia legítimas, así como la estructura de las conclusiones que ellas permiten deducir de cada clase tipificable de premisas. En este aspecto esencial el programa de Hilbert se apoya en la formalización de la lógica y las matemáticas emprendida por el propio Frege y perfeccionada por Russell y Whitehead. En los Capítulos 2.2–2.4 hablaré de las motivaciones y principales logros de esa empresa. Pero antes debo redondear este capítulo introductorio con algunas indicaciones generales sobre la función que desempeñan en el programa de Hilbert.

En *Principia Mathematica* (1910–13), Whitehead y Russell mostraron que todas las proposiciones clásicas de la aritmética y el análisis pueden enunciarse en la escritura ideográfica, de sintaxis simple y precisa, desarrollada por ellos. En su sistema lógico-matemático las inferencias desde premisas formuladas en esa escritura se ajustan a unos pocos patrones fácilmente reconocibles. Utilizando esos patrones, de un par de premisas contradicto-

rias puede deducirse *cualquier* conclusión.[12] El programa de Hilbert se basa en esto. Hilbert piensa que una vez que los axiomas cuya consistencia interesa demostrar se han enunciado en la escritura canónica podemos desatender por completo lo que se quería decir con ellos (olvidar, por ejemplo, que intentábamos referirnos al infinito actual). La única tarea pendiente consiste en probar que hay por lo menos una oración que no se puede deducir de los axiomas. Específicamente, es preciso demostrar que es imposible que cierta fórmula de la escritura canónica—por ejemplo, un enunciado de la igualdad ‘0 = 1’— ocupe el último lugar en una lista de tales fórmulas cada una de las cuales es un axioma o se deduce de una o más de sus predecesoras en la lista conforme a uno de los patrones de inferencia del sistema. Tal demostración, claro está, debe efectuarse de manera que no deje lugar a dudas y que no invite, a su vez, a probar la consistencia de los recursos utilizados en ella. Para ello es necesario y suficiente, según Hilbert, que la demostración requerida se conduzca enteramente en el ámbito de lo que él llama la “postura finita” (*finite Einstellung*) o el “punto de vista finito” (*finiter Standpunkt*).

De este modo, Hilbert, sin renunciar al transfinito cantoriano, trata de hacerlo inofensivo reduciéndolo a ser nada más que el referente ideal —y olvidable— de ciertos símbolos que el matemático puede y debe utilizar como si no significaran nada. Ello implica reemplazar lo que Hilbert llama el “razonamiento sustantivo” (*inhaltliches Schließen*), expresado en alemán o castellano, que se hace pensando en aquello sobre lo cual se discurre, por un tipo de razonamiento puramente formal, que, sin pensar, manipula filas de signos de la escritura ideográfica conforme a reglas que conciernen únicamente a su aspecto y no a su significado.[13] Esto no quiere decir, sin embargo, que Hilbert auspicie una matemática no-pensante. A la matemática “clásica”, que

[12] Sea $\neg\alpha$ la negación de una oración $\alpha$. Sea $\beta$ otra oración cualquiera. De la premisa $\alpha$ es lícito inferir $(\beta \to \alpha)$. De las premisas $(\beta \to \alpha)$ y $\neg\alpha$ es lícito inferir $\beta$. Por lo tanto, de las premisas $\alpha$ y $\neg\alpha$, puede inferirse la conclusión $\beta$.

[13] “Obtenemos así, en lugar de la ciencia matemática sustantiva (*inhaltlich*), trasmitida en el lenguaje corriente, un caudal de fórmulas con signos matemáticos y lógicos, que se ponen en fila según ciertas reglas. A los axiomas matemáticos corresponden algunas de las fórmulas y a la inferencia sustantiva (*inhaltliches Schließen*) corresponden las reglas conforme a las cuales las fórmulas se suceden unas a otras. La inferencia sustantiva es reemplazada, pues, por un operar extrínseco según reglas . . . ” (Hilbert, 1925; GA, III, 177). El vocablo alemán *inhaltlich* que traduzco por ‘sustantivo, a’, se deriva de *Inhalt*, ‘contenido’ y significa literalmente ‘propio del contenido’, ‘relativo al conteni-

se trata de hacer inmune a la crítica aunque sea al precio de exonerarla del pensamiento, viene ahora a agregarse una "meta-matemática": la Teoría de la Prueba (*Beweistheorie*), que estudia sistemáticamente el razonamiento formal y servirá para establecer que éste no podrá nunca generar una contradicción dentro de la matemática formalizada. La Teoría de la Prueba razona "sustantivamente", pensando en lo que dice; pero, como trata sólo de signos que son marcas fácilmente reconocibles de tinta o de tiza y de reglas simples para combinar esas marcas en filas y en columnas, se mantiene, según Hilbert, dentro de la "postura finita" y no corre el riesgo de contradecirse.[14]

Hilbert nunca hizo un inventario de los recursos accesibles a la "postura finita" ni los caracterizó mediante una definición precisa. Pero los ejemplos que ofrece en los artículos en que anuncia y defiende el programa, la exigencia de que la prueba de consistencia sea inequívoca e incuestionable, y el mismo adjetivo 'finito' empleado para describirlos dan, creo, una idea bastante clara de la índole y alcance de esos recursos. Parece obvio, por ejemplo, que el matemático en "postura finita" no podrá contemplar la existencia de un conjunto cerrado de objetos actualmente infinito, sino a lo sumo secuencias abiertas, potencialmente infinitas. Al razonar sobre éstas, se abstendrá de invocar el Principio del Tercero Excluido. Como anota von Neumann: "El demostrar sustantivo debe desarrollarse en completo acuerdo con la lógica intuicionista de Brouwer y Weyl. La Teoría de la Prueba debe edificar, por decirlo así, la matemática clásica sobre una base intuicionista y de este

---

do'. Por eso el término hilbertiano *inhaltlich* se ha solido traducir al inglés por 'contentual', un neologismo que es feo en ese idioma y sería prácticamente ininteligible en el nuestro. Me parece en cambio que el contraste que hacemos en castellano corriente (no filosófico) entre la "mera forma" y la "sustancia" de un asunto justifica mi traducción de *inhaltlich*.

14 Von Neumann (1927, pp. 1–2) describe estos dos tipos de razonamiento con insuperable claridad:

> Hay que distinguir tajantemente entre dos modos diferentes de "demostrar": El demostrar formalista ("matemático") dentro del sistema formal y el demostrar sustantivo ("metamatemático") concerniente al sistema. Mientras aquél es un juego lógico definido arbitrariamente (aunque tiene que ser, por cierto, en buena medida análogo a la matemática clásica), éste consiste en un encadenamiento de intuiciones sustantivas (*inhaltlicher Einsichten*) inmediatamente evidentes.

modo reducir al absurdo el intuicionismo estricto" (1927, p. 2).[15]

La insistencia en que las teorías matemáticas bajo estudio se formulen en una escritura canónica similar la adoptada en *Principia Mathematica* obedece, sobre todo, a que entonces —como ha señalado Kreisel— la cuestión de su consistencia se torna "finita" en el sentido preciso siguiente: "Dadas las reglas formales y, por lo tanto, en particular, la lista (finita) de símbolos a utilizar, uno puede decidir con respecto a cualquier secuencia de 'palabras' (= fórmulas) si está construida de acuerdo con las reglas y si contiene o no palabras de la forma $a$ y $\neg a$, donde el símbolo $\neg$ se usa para expresar la *negación*" (Kreisel 1976, p. 94). Este es el objetivo de las tres primeras tareas que von Neumann asigna a la Teoría de la Prueba en la ponencia que leyó como portavoz del grupo de Hilbert en el simposio sobre fundamentos de la matemática celebrado en Königsberg en septiembre de 1930 (en que también hablaron Heyting, por la escuela de Brouwer, y Carnap, por el logicismo *à la* Frege-Russell favorecido en el Círculo de Viena):

1. Enumerar todos los símbolos empleados en la matemática y la lógica.

2. Caracterizar inequívocamente todas las combinaciones de dichos símbolos que representan oraciones dotadas de sentido según la matemática clásica. Estas combinaciones se llaman "fórmulas". (Adviértase que una oración dotada de sentido no tiene que ser verdadera. '1 + 1 = 1' es una oración dotada de sentido, pero falsa).

[15] La última cita sigue inmediatamente al texto traducido en la nota 14. Justifica el uso común en la escuela de Hilbert del epíteto 'intuicionista' para calificar los métodos de razonamiento "sustantivo" admitidos por ella. Herbrand (1931; EL, p. 225, n. 3) define dicho epíteto con admirable claridad:

> Entendemos por razonamiento intuicionista un razonamiento que satisfaga las siguientes condiciones: en él se considera sólo un número finito determinado de objetos y de funciones; éstas están bien definidas, de modo que su definición permita calcular unívocamente su valor; nunca se afirma la existencia de un objeto sin dar el medio de construirlo; nunca se considera el conjunto de todos los objetos $x$ de una colección infinita; y cuando se dice que un razonamiento (o un teorema) es verdadero para todo $x$, esto significa que para cada $x$ tomado particularmente se puede repetir el razonamiento general en cuestión, el cual no debe considerarse sino como el prototipo de estos razonamientos particulares.
>
> (Herbrand, EL, p. 225 n.3)

3. Determinar un procedimiento que permita construir sucesivamente todas las fórmulas que corresponden a oraciones “demostrables” de la matemática clásica. Tales construcciones se llaman “pruebas”.

Como estas tres tareas ya fueron ejecutadas casi a cabalidad por Whitehead y Russell antes de que Hilbert siquiera concibiese su programa, sólo queda por resolver una cuarta y última tarea, que von Neumann describe aproximadamente en estos términos:

4. Mostrar “de un modo combinatorio finitista” que las fórmulas correspondientes a oraciones de la matemática clásica que pueden verificarse por métodos aritméticos finitistas pueden “probarse” —esto es, construirse mediante el procedimiento mencionado en el punto 3— sí y sólo si la verificación de la oración correspondiente muestra que ésta es verdadera.

Con el cumplimiento de esta tarea quedará establecida la consistencia de la matemática formalizada (puesto que si fuese inconsistente se podría “probar” en ella la fórmula representativa de una igualdad numérica falsa). Pero se establecerá además que la fórmula representativa de *cualquier* verdad aritmética elemental puede “probarse” en la matemática formalizada. En otras palabras, el programa de Hilbert, según lo explica von Neumann en 1930, proponía demostrar, no sólo que la matemática formalizada es consistente, sino además que constituye una representación formal *completa* de la aritmética elemental.[16] Esta doble demostración tiene que llevarse a cabo ente-

16 Hilbert (1930, p. 6) menciona “la completud (*Vollständigkeit*) del sistema de axiomas de la aritmética (*Zahlentheorie*) así como del análisis” entre los “problemas de fundamentación de las matemáticas” aún pendientes. Aunque dicha completud “comúnmente se admite, la consideración usual mediante la cual se muestra que cualquier par de realizaciones del sistema de axiomas de la aritmética (respectivamente, del análisis) tienen que ser isomórficas, no satisface las exigencias del rigor finito”. Hilbert propone dar en cambio —“por de pronto para la aritmética, cuyo campo se deja delimitar con precisión”— un sustituto finitista de la prueba habitual de isomorfismo, demostrando que “si es posible probar la consistencia de una oración S con los axiomas de la aritmética, no es posible probar la consistencia de ¬S (la contradictoria de S) con esos axiomas”. Adviértase que el trabajo en que Hilbert propone este problema no fue incluido en sus *Gesammelte Abhandlungen* (editados después de la publicación de los teoremas de incompletud de Gödel).

ramente con los recursos del *análisis combinatorio finito*. Por una curiosa coincidencia, Gödel anunció en la misma reunión de Königsberg sus célebres teoremas de incompletud (Gödel 1930a, 1931). Los estudiaremos en el Capítulo 2.10. Por ahora, baste indicar, *grosso modo*, que en virtud de ellos, si $\mathcal{T}$ es una formalización de la aritmética elemental ajustada a las condiciones 1, 2 y 3 no es posible que $\mathcal{T}$ sea a la vez consistente y completa. Además, aunque en dicha formalización $\mathcal{T}$ habrá más de una fórmula apta para representar la oración '$\mathcal{T}$ es consistente', ninguna de ellas se puede "probar" en $\mathcal{T}$ a menos que $\mathcal{T}$ sea inconsistente. Como los razonamientos del análisis combinatorio finito ciertamente se dejarían representar mediante "pruebas" en una tal formalización de la aritmética, sería imposible ejecutar la tarea cuarta y final del programa de Hilbert "de modo combinatorio finitista". Pero Hilbert nunca había explicado con tanta precisión como von Neumann el significado de su "postura finita" y gracias a eso pudo eventualmente "ampliar el marco metódico" de su programa a la luz de los hallazgos de Gödel, sin admitir su fracaso.[17]

[17] Véase la Sección 2.9.4. También el Capitulo 2.12, primer párrafo y nota 1, donde me refiero a Hilbert y Bernays 1939, § 5, "Der Anlaß zur Erweiterung des methodischen Rahmens der Beweistheorie" ("El motivo para ampliar el marco metódico de la Teoría de la Prueba"). El tomo I de la misma obra contiene la siguiente explicación, elocuente pero no muy precisa, sobre el significado del adjetivo 'finit' (que generalmente traduzco 'finitista'):

> Con la palabra "finit" queremos siempre expresar que la reflexión, aseveración o definición así calificada se ciñe a los límites de la representabilidad de objetos y de la ejecutabilidad de procesos y por lo tanto se realiza en el marco de una consideración concreta.
>
> (Hilbert y Bernays 1934, p. 32)

Donde escribo 'representabilidad' y 'ejecutabilidad', los autores dicen "grundsätzliche Vorstellbarkeit" ("representabilidad en principio") y "grundsätzliche Ausführbarkeit" ("ejecutabilidad en principio"), pero me ha parecido inútil agregar una indicación que lastra la frase y que cualquier lector de inteligencia mediterránea sobreentiende.

## 2.2 ESCRITURA CONCEPTUAL

Leibniz soñó con "una lengua o escritura universal" en que "los caracteres y las palabras dirigirían la razón y los errores que no fuesen errores de hecho no serían sino errores de cálculo" (GP, III, 605). En dicho lenguaje artificial, toda falacia del razonamiento estaría vedada por la sintaxis, de modo que

> un *sofisma*, expresado en tal nueva forma de escritura, no sería en verdad más que un *solecismo* o *barbarismo*, fácil de refutar aduciendo las leyes de esa gramática filosófica. Así, cuando surgieran desacuerdos entre dos filósofos, no haría falta más discusión que entre un par de calculistas. Bastaría que, pluma en mano, —en presencia, si se quiere, de un amigo— se sentasen a sus ábacos, diciéndose mutuamente: *calculemos*.
>
> (Leibniz, GP, VII, 200)[1]

Para lograr este resultado, Leibniz propuso extender a la lógica los métodos algebraicos utilizados con tanto éxito en aritmética y geometría. Entre sus papeles quedaron —inéditos hasta el siglo XIX— algunos ensayos rudimentarios de lógica algebraica. Pero la idea de un álgebra de la lógica sólo toma vuelo con los libros de Boole (1847, 1854).[2] La tradición booleana, cultivada en Inglaterra por Jevons y en América por Peirce, fue recogida y promovida en Alemania por Ernst Schröder (1877, VAL: 1890–1905). Para el filósofo norteamericano C. I. Lewis —que publica un *Panorama de la lógica simbólica* en 1918— "el álgebra clásica de la lógica" lleva los apellidos de Boole y Schröder. En cambio, Hilbert, que sin duda los conocía, no suele nombrarlos entre los precursores de su programa, cuya iniciación atri-

1 Véase asímismo Leibniz, GP, VII, 204–207 (traducción castellana en Leibniz, EF, pp. 188–193).

2 Los aportes más interesantes de Leibniz son quizás los dos "especímenes" en GP, VII, 228–247 (traducción inglesa en Leibniz, LP, pp. 122–144). En el siglo XVIII hubo algunos intentos aislados por desarrollar un cálculo lógico; al respecto, véase C. I. Lewis 1918, pp. 18–51, quien destaca sobre todo la obra de J. H. Lambert.

buye, como hemos visto, a Gottlob Frege —a quien Lewis dedica menos de una página de su *Panorama* (1918, pp. 114s.)— y a sus continuadores Russell y Whitehead.

Varias razones justifican, a mi modo de ver, esta preferencia de Hilbert por Frege. En primer lugar, Frege creó su propia lengua artificial (*Escritura conceptual*, 1879) con el propósito específico de expresar en ella razonamientos matemáticos y utilizarla en la fundamentación estrictamente lógica de la aritmética. En cierto modo, pues, el proyecto de Frege anticipa directamente el programa de Hilbert. Además —como reconoce el propio Lewis— Frege entendió antes que nadie los requisitos que ha de cumplir la representación escrita de un razonamiento deductivo para que la corrección o incorrección del mismo pueda controlarse mediante un cálculo. Pero, sobre todo, debemos a Frege el análisis moderno de los predicados relacionales y la generalización universal, que hizo posible entender muchas demonstraciones matemáticas como casos de simple inferencia lógica.[3]

Frege abraza sin reservas la clasificación kantiana de las verdades en analíticas y sintéticas. Mejorando a Kant, caracteriza ambas clases de modo que constituyan una genuina partición: verdades analíticas son las que se deducen de leyes lógicas y meras definiciones; cualquier verdad que no es analítica es sintética.[4] La meta de sus investigaciones lógico-matemáticas es demostrar que las verdades de la aritmética y del análisis son, en este sentido,

[3] Conviene, sí, recordar que también Peirce (1870, 1880, 1882, 1883) desarrolló por esos mismos años —independientemente de Frege— la lógica moderna de la generalización y de los predicados poliádicos. Por otra parte, como ha mostrado Goldfarb (1979), la generalización vino a entenderse cabalmente como la entendemos ahora sólo después de 1920, precisamente en la escuela de Hilbert.

[4] Para Kant una aseveración es analítica si el predicado está contenido en el concepto del sujeto y es sintética si el predicado no está contenido en el concepto del sujeto. Las aseveraciones disyuntivas y condicionales, que no se dejan analizar en sujeto y predicado, no caen, pues, en ninguna de estas dos clases, según Kant. Frege tuvo que proponer una caracterización diferente porque en un comienzo rechazó el análisis tradicional de las aseveraciones en sujeto y predicado (Frege 1879, p. 3). Su caracterización aventaja también a la kantiana en cuanto permite clasificar sin dificultada como analíticas a las aseveraciones que expresan leyes lógicas. En efecto, como indiqué en el Capitulo 2.1, decimos que una aseveración $\alpha$ se deduce de un conjunto de aseveraciones $\mathscr{S}$ si $\alpha$ ocupa el último lugar en una lista de aseveraciones cada una de las cuales pertenece a $\mathscr{S}$ o se infiere directamente de una o más de sus predecesoras en la lista. Por lo tanto, si $\mathscr{S}$ es un conjunto de aseveraciones que expresan leyes lógicas y $\alpha$ es

analíticas. Para presentar la demostración de tal modo que no dejase lugar a dudas, Frege inventó su "Begriffschrift" ("escritura conceptual"), un lenguaje artificial especialmente diseñado para la formulación perspicua de inferencias deductivas, que designaré aquí con la sigla BS. BS es una versión de lo que hoy llamamos un cálculo predicativo. Frege admite la posibilidad de hacer generalizaciones referentes a propiedades y relaciones, pero no la reglamenta expresamente. Si BS* es lo que resta de BS cuando se elimina esta posibilidad y sólo se permiten generalizaciones referentes a objetos individuales, tenemos que BS* equivale a nuestro cálculo predicativo de primer orden.[5]

Aunque Frege maneja el lenguaje BS con impecable coherencia, sus explicaciones gramaticales no siempre alcanzan la claridad y precisión a que nos han habituado sus seguidores, los modernos tratadistas de lógica. No llega a dar un inventario formal de las letras que pueden usarse como variables y constantes de BS, ni una codificación exhaustiva del modo de utilizarlas.[6] Además, en su presentación original, la elucidación de las nociones decisivas de 'contenido' y de 'función' deja mucho que desear. Frege corrige este defecto en la presentación final de BS en el tomo I de *Leyes fundamentales de la aritmética* (1893), aprovechando las grandes innovaciones ontológicas y semánticas introducidas en "Función y concepto" (1891), "Sobre sentido y denotación" (1892) y "Sobre concepto y objeto" (1892a). No puedo entrar

---

una de ellas, diremos que $\alpha$ se deduce de $\mathscr{S}$, por cuanto $\alpha$ ocupa ciertamente el último lugar de la lista cuyo único miembro es $\alpha$ y, por hipótesis, $\alpha \in \mathscr{S}$.

5 Antes de leer el resto de este capítulo —y los siguientes— el lector que no haya estudiado algo de lógica moderna debiera darle una ojeada al Apéndice IX sobre el cálculo predicativo. Como la exposición es bastante concisa, no creo que le aproveche a quien no tenga ya cierta práctica en la lectura de obras matemáticas. El lector que no la tenga debe leer uno o dos manuales de lógica. Recomiendo a Jeffrey 1981 y Mates 1970, cuyas virtudes se complementan.

6 Ello genera sorpresas. Así, a la luz de los ejemplos de Frege 1879, p. 51, y de una lectura desaprensiva de los §§ 9–12, el lector pensará ingenuamente que, en la expresiones $f(b)$, $g(b)$, $h(b)$ de la Proposición 60 (p. 52), la letra $b$ es una constante o variable individual, y las letras $f$, $g$ y $h$ son predicados. Pero en la p. 69 se nos manda sustituir, en la Proposición 60, la letra $b$ por la variable predicativa $\mathfrak{F}$ y las expresiones de la forma $f(\Gamma)$ por expresiones de la forma $\Gamma(y)$, de modo que $f(b)$ ha de reemplazarse con $\mathfrak{F}(y)$.

aquí a comentar esas innovaciones,[7] aunque daré al respecto una indicación somera que ayude a entender la gramática de BS.

Los signos de BS son de dos clases: *letras*, "con las que uno puede representarse una variedad de cosas", y *figuras lineales* "que tienen un sentido completamente determinado". Las letras, de varios tipos, "deben servir principalemente para expresar la *generalidad*" (1879, p. 1).[8] Como dije, Frege no explica metódicamente cómo se usan los distintos tipos de letras. Sólo el empleo de las letras góticas $\mathfrak{a}$, $\mathfrak{F}$, . . . está cuidadosamente reglamentado: una letra gótica sólo puede figurar en una expresión si va precedida de un cuantificador que la contiene; en otras palabras, en BS una letra gótica es siempre una variable ligada (individual, si es minúscula; predicativa, si es mayúscula). Entiendo, en cambio, que las mayúsculas del alfabeto griego, Α, Β, Γ, . . . funcionan como constantes, que algunas veces significan una oración completa (así, la A y la B en Frege 1879, p. 22), otras veces un predicado determinado (vgr. la Φ en esa página).[9] Las minúsculas del alfabeto griego, α, β, γ, . . . se usan en *Escritura conceptual* como índices —esto es, variables ligadas— en ciertas expresiones simbólicas definidas (1879, §§ 24, 26) que se dejan eliminar; también figuran en las indicaciones marginales sobre sustitución de expresiones, pero en este uso no pertenecen a BS, sino a la lengua viva —el alemán— en que se dan esas indicaciones: son pues lo que hoy llamamos 'variables metalingüísticas'. En *Leyes fundamentales de la aritmética* las minúsculas griegas tienen también otros usos que no es preciso detallar aquí. Por último, las cursivas latinas minúsculas se emplean como variables libres: *a*, *b*, *c*, *m* y *x* funcionan a veces como variables proposicionales (Proposiciones 1–51), a veces como variables individuales (Proposiciones 52–62, 64–65), a veces como ambas cosas (Proposiciones 63, 67 y 68); en cambio, *f*, *g*, *h* y la mayúscula *F* son variables

[7] Los tres artículos mencionados están bien traducidos al castellano en Frege, LS.

[8] Frege dice que esta clasificación de los signos y sus funciones se inspira en el ejemplo de la matemática que usa letras *a*, *b*, *x*, *y*, . . . con significado variable e ideogramas como +, ×, √ con un significado fijo. Pero los matemáticos, según él, no son enteramente consecuentes, puesto que usan combinaciones de letras, como log, sen, Lim, como ideogramas (1879, p. 1 n.).

[9] Frege 1879, p. 2n., dice que las mayúsculas griegas son "abreviaturas" a las que el lector ha de asignar cualquier sentido apropiado, cuando no estén expresamente definidas por el autor.

predicativas (monádicas o diádicas).

Las figuras lineales en BS se componen de los signos siguientes: ı , ── , ─┬─ , ┬└ , ◡ y =. Como luego veremos, = (en 1879, ≡) es el signo de identidad; ◡ es el signo de cuantificación universal (la variable ligada anida en la cucharita), y ─┬─ y ┬└ corresponden respectivamente a nuestros signos de negación (¬) e implicación (→). Pero los signos ı y ── no tienen su homólogo en el cálculo predicativo actual. Frege los explica así:

> Una aseveración (*Urtheil*) se expresará siempre mediante el signo
>
> ├──
>
> situado a la izquierda del signo o combinación de signos que indica el contenido de la aseveración. Al *omitir* el breve trazo vertical en el extremo izquierdo del horizontal, la aseveración se convierte en una *mera combinación de representaciones* (*blosse Vorstellungsverbindung*), de la cual el escribiente no expresa si le atribuye verdad o no. Por ejemplo, si
>
> ├── A
>
> significa la aseveración "los polos magnéticos de distinto nombre se atraen", entonces
>
> ── A
>
> no expresa la aseveración, sino debe evocar en el lector únicamente la representación de la atracción mutua de los polos magnéticos de distinto nombre, tal vez para derivar de ella consecuencias y examinar a la luz de éstas si el pensamiento (*Gedanke*) es correcto. En este caso, *parafraseamos* la expresión con las palabras "*la circunstancia de que*" o "*la proposición:*" („*der Satz, daß*"). [ . . . ] *El trazo horizontal* con que se forma el signo ├── *reúne los signos siguientes en un todo, y a este todo se refiere la afirmación expresada por el trazo vertical en el extremo izquierdo del horizontal*.
>
> (Frege 1879, pp. 1–2)[10]

10 Nótese que Frege identifica aquí la *representación* evocada por los signos —que puede variar mucho de lector en lector— con el *pensamiento* expresado mediante ellos —que ha ser el mismo para todos si la escritura sirve como medio de comunicación—; más tarde distinguirá enfáticamente estas dos cosas. Más grave me parece la confusión si-

Frege recuerda enseguida que no todo contenido expresable verbalmente puede ser materia de una aseveración. Puedo aseverar: ‘la casa de Juan es de madera’; pero no puedo aseverar: ‘casa’. Cabe, pues, distinguir entre *contenidos aseverables* y *no aseverables* (“*beurtheilbare* und *unbeurtheilbare* Inhalte”). Llamemos *simple* a una aseveración cuyo contenido aseverable no incluye partes (distintas de él) que a su vez son contenidos aseverables. La lógica tradicional analizaba el contenido aseverable de cualquier aseveración simple en dos componentes no aseverables: *sujeto* y *predicado*. Frege rechaza este análisis (1879, p. 2). Hace ver, en cambio, que un mismo contenido, aseverable o no, puede analizarse de diversas maneras, fijando una parte de la expresión que lo enuncia y tratando el resto como variable. Así, en la expresión ‘Juan admira a Juanita’, si fijo ‘admira a Juanita’ y reemplazo ‘Juan’ por ‘Pedro’, ‘Diego’, etc. efectúo un análisis similar al tradicional; si fijo ‘Juan admira a’ y reemplazo ‘Juanita’ por ‘Isabel’, ‘Beatriz’, etc., llego a un análisis correspondiente al tradicional de ‘Juanita es admirada por Juan’, que es el mismo contenido aseverable, expresado en voz pasiva; pero si sólo fijo ‘admira a’ y reemplazo el par ⟨‘Juan’, ‘Juanita’⟩ por ⟨‘Pedro’, ‘Isabel’⟩, ⟨‘Diego’, ‘Beatriz’⟩, ⟨‘Diego’, ‘Juanita’⟩, etc., o si fijo ⟨‘Juan’, ‘Juanita’⟩ y reemplazo ‘admira a’ por ‘saluda a’, ‘sonríe a’, ‘detesta a’, etc. obtengo descomposiciones del contenido aseverable que no tienen precedente en la lógica tradicional. Los contenidos no aseverables admiten el mismo tratamiento: por ejemplo, en la frase ‘la casa de Juan’, puedo fijar ‘la casa’ y reemplazar ‘de Juan’ por ‘de Diego’, ‘de Pedro’, etc. o fijar ‘de Juan’ y reemplazar ‘la casa’ por ‘la mujer’, ‘la profesión’, ‘el país’, etc., o fijar ‘de’ y tratar el par ⟨‘la casa’, ‘Juan’⟩ como variable.[11] En *Escritura conceptual*, Frege

guiente: Frege sugiere que uno puede derivar consecuencias de un pensamiento no aseverado (y examinar la corrección de éste a la luz de aquéllas), pero según las reglas de inferencia de BS, las premisas de una deducción tienen que ser aseveraciones (esto es, expresiones de la forma ⊢A, nunca de la forma —A).

11 La simetría que Frege evidentemente postula entre estos diversos modos de análisis hace comprensibles las sustituciones descritas en la nota 6. La libertad con que Frege está dispuesto a ver en una expresión como ‘León ríe’, tanto una expresión del valor de la función *ríe* para el argumento *León* (reemplazable por *Pedro*, *Juan*, *Diego*, etc.) como una expresión del valor de la función *León* para el argumento *ríe* (reemplazable por ‘*llora*’, ‘*ruge*’, ‘*muerde*’, etc.) es familiar en matemáticas. He aquí un caso análo-

—inspirándose en la terminología matemática— llama *función* al componente fijo, *argumento* al componente reputado variable de una expresión representativa de un dado contenido, aseverable o no.[12] Pero en *Leyes fundamentales de la aritmética* rechaza sin más el error de quienes consideran a la función "como una *expresión,* una combinación de signos, y no como lo designado por ellos" (1893, p. 5).

Hacia 1891 Frege llegó a la conclusión de que en el significado de las expresiones lingüísticas hay que distinguir dos aspectos complementarios pero bien diferentes: su *sentido* (*Sinn*) y lo que llamaré su *denotación* (Frege lo llama *Bedeutung*; literalmente, 'significado').[13] Las expresiones '$7^2$', '100 – 51' y '(11 × 2) + (9 × 3)' denotan lo mismo, a saber, el número 49, pero no lo presentan del mismo modo, así que no tienen el mismo sentido. Una igualdad, como $13^5 = 371.293$, sería falsa si el miembro de la izquierda no denotase el mismo objeto que el miembro de la derecha, pero sería banal si ambos miembros tuviesen el mismo sentido, como en $13^5 = 13^5$. Por ese mismo tiempo, Frege (1891, 1892a) revisa a fondo su distingo entre función y argumento. Observa que hay dos clases de expresiones: expresiones "saturadas" (*gesättigt*), cuyo sentido está completo o "cerrado" (*abgeschlossen*), y ex-

go, tomado del álgebra lineal: Si $\mathcal{V}$ es un espacio vectorial sobre el cuerpo $\mathbb{R}$, las aplicaciones lineales de $\mathcal{V}$ en $\mathbb{R}$ forman otro espacio vectorial $\mathcal{V}^*$ sobre $\mathbb{R}$. Las aplicaciones lineales de $\mathcal{V}^*$ en $\mathbb{R}$ se identifican con los vectores de $\mathcal{V}$ mediante la estipulación siguiente: para cada $\mathbf{r} \in \mathcal{V}$, $\mathbf{r}: \mathcal{V}^* \to \mathbb{R}$ es la aplicación (demostrablemente lineal) que en cada $\xi \in \mathcal{V}^*$ toma precisamente el valor que $\xi: \mathcal{V} \to \mathbb{R}$ toma en $\mathbf{r}$. Se acostumbra designar dicho valor con $\langle \xi | \mathbf{r} \rangle$, una expresión que puede leerse de dos maneras, a saber (i) 'el valor $\xi(\mathbf{r})$ de la función lineal $\xi: \mathcal{V} \to \mathbb{R}$ en el argumento $\mathbf{r} \in \mathcal{V}$' y (ii) 'el valor $\mathbf{r}(\xi)$ de la función lineal $\mathbf{r}: \mathcal{V}^* \to \mathbb{R}$ en el argumento $\xi \in \mathcal{V}^*$'.

12 Cito la definición pertinente:

> Si en una *expresión* (*Ausdruck*) —cuyo contenido no tiene que ser aseverable— figura en una o más posiciones un *signo* (*Zeichen*) simple o compuesto y pensamos que este signo es reemplazable en algunas de esas posiciones o en todas ellas por otra cosa (*durch Anderes*), pero que sea la misma en cada caso, llamamos *función* a la parte de la expresión que se presenta en este respecto como invariable y a la parte reemplazable la llamamos su *argumento*.
> (Frege 1879, p. 18; énfasis mío; el pasaje entero va en cursiva en el original)

13 El distingo entre *Sinn* y *Bedeutung* es el tema de Frege 1892, pero ya aparece mencionado en Frege 1891, p. 14.

presiones “insaturadas” (*ungesättigt*), que entendemos como provistas de uno o más blancos a llenar con expresiones saturadas. En los ejemplos del párrafo anterior, ‘Juan’, ‘la casa de Juan’, ‘Juan admira a Juanita’ son expresiones saturadas; en cambio, ‘la casa de …’, ‘… de …’, ‘… admira a Juanita’, ‘Juan admira a …’, ‘… admira a …’ son expresiones insaturadas (en las que ahora he marcado blancos donde antes los subentendíamos). Según Frege, la diferencia entre expresiones saturadas e insaturadas refleja la división de lo real en dos categorías irreductibles: *objetos* (*Gegenstände*) y *funciones* (*Funktionen*). Aquéllos son, en cierto modo, completos y se los denota con expresiones saturadas; éstas son incompletas y se las designa mediante expresiones insaturadas. Por ejemplo, ‘3’ y ‘7’ denotan objetos; ‘… + …’ designa una función. La expresión saturada ‘3 + 7’ denota el objeto que es el valor de la función …+… en el par de argumentos ⟨3,7⟩; el mismo objeto es denotado también por la expresión saturada ‘10’.

A la luz de estas ideas nuevas, es claro que lo que Frege llamaba en 1879 “contenido aseverable” reúne —y confunde— el sentido y la denotación de la expresión utilizada para hacer una aseveración (1893, p. 9 n. 2). Toda expresión utilizable para ello, es decir, toda oración declarativa, es por cierto una expresión saturada. Su sentido es lo que Frege llama un *pensamiento* (*Gedanke*); su denotación, uno de los dos objetos que llama *valores veritativos* (*Wahrheitswerte*): *lo verdadero* (*das Wahre*) o *lo falso*. Una oración declarativa simple —no formada por combinación o transformación de otras oraciones declarativas— constará en todo caso de una o más expresiones saturadas que denotan objetos, insertadas en los huecos de una expresión insaturada que denota un concepto. Un concepto es pues una función que admite como argumento cualquier $n$-tuplo de objetos (para un entero positivo fijo $n$, característico de ese concepto), pero cuyo valor para un dado argumento sólo puede ser uno de los dos valores veritativos, a saber, lo verdadero, si el $n$-tuplo en cuestión cae bajo el concepto, y lo falso, si dicho $n$-tuplo no cae bajo el concepto.

El distingo entre objetos “llenos” y funciones “huecas” evoca la clásica división de lo real en particulares y universales, combinada con la tesis occamista de que sólo los particulares existen plenamente, mientras que los universales sólo pueden actualizarse en aquellos. El distingo entre dos clases de expresiones correspondientes a las dos categorías de entes da nueva vida al análisis de las oraciones en sujeto y predicado, rechazado por Frege en

1879. Pero este análisis rebasa ahora los estrechos límites a que lo confinaba la tradición. Según ella, cada aseveración simple se refería a un solo objeto, el sujeto, denotado por una expresión saturada, y le atribuía una propiedad, designada con una expresión insaturada, el predicado. En cambio, como Frege entiende que las expresiones insaturadas denotan *funciones* —en una acepción del vocablo inspirada en la terminología matemática—, puede con toda naturalidad admitir que una aseveración simple se refiera a un $n$-tuplo de objetos ($n = 1, 2, \ldots$) y les atribuya una relación (si $n > 1$). Por otra parte, a la luz del análisis fregeano, es claro que un predicado (expresión insaturada) no puede hacer las veces de sujeto (expresión saturada). Así, si '$Hp$' significa 'Píndaro es hombre' y '$Gp$' significa 'Píndaro es griego', no es lícito poner en la primera oración '$G$' en el lugar de '$p$' para significar (i) 'Todo griego es hombre' o (ii) 'Algún griego es hombre'. En estas oraciones del lenguaje corriente, el sujeto gramatical —'todo griego', 'algún griego'— no denota un objeto al cual se atribuye el predicado sino que circunscribe la clase a que ese objeto pertenece (mediante un predicado común a todos los miembros de esa clase). Así, según Frege, lo que dicen las oraciones citadas queda mejor expresado por (i) 'Si algo es griego, también es hombre' y (ii) 'Hay algo que es griego y también hombre' (en el simbolismo estándar del cálculo predicativo: (i) $\forall x(Gx \to Hx)$; (ii) $\exists x(Gx \wedge Hx)$). Gracias a estas innovaciones, Frege pudo poner de manifiesto, como dije, la índole puramente lógica de muchas demostraciones matemáticas.[14]

Ahora puedo dar con Frege una definición más clara y satisfactoria de la línea — y de los cuatro signos de BS que arriba dejamos sin definir (cf. Frege 1893, pp. 9–12). Usaré la letra negrita **v** para nombrar lo verdadero y la **f** para nombrar lo falso. La línea — designa la función cuyo valor es **v** en el argumento **v** y cuyo valor es **f** en cualquier otro argumento.[15] El signo —┬—

[14] Por ejemplo, de las premisas "√2 es la raíz cuadrada positiva de dos" (que representaré con '$R_2rd \wedge r > 0$'), "Cuatro es el cuadrado de dos" ('$Qcd$') y "Toda raíz cuadrada de un número es una raíz cuarta del cuadrado de ese número" ('$\forall x \forall y(Qyx \to \forall z(R_2zx \to R_4zy))$') se deduce fácilmente que "√2 es una raíz cuarta de cuatro" ('$R_4rc$'), una conclusión obvia que sin embargo es inaccesible a la lógica aristotélica.

[15] Así pues, la función —, restringida a valores veritativos, es precisamente la aplicación $x \mapsto x$ (—$\Gamma = \mathbf{v}$ si $\Gamma = \mathbf{v}$ y —$\Gamma = \mathbf{f}$ si $\Gamma = \mathbf{f}$), lo cual explica que no haya una signo para ella en el simbolismo lógico actual: normalmente sería superfluo. Pero Frege no deslinda explícitamente el dominio de sus funciones, a las que trata como si cada una

designa la función cuyo valor es **f** en cada argumento *x* tal que —*x* es **v** y cuyo valor es **v** en cada argumento *x* tal que —*x* es **f** (nuestro signo de negación ¬ se inspira probablemente en este signo fregeano).[16] La expresión ‘$\smile\!\!\!\!^{\mathfrak{a}}$— $f(\mathfrak{a})$’ denota **v** si y sólo si la función *f*( ) toma el valor **v** en todo argumento. La expresión ‘$\smile\!\!\!\!^{\mathfrak{F}}$— $\mathfrak{F}$(A)’ denota **v** si y sólo toda función toma el valor **v** en el argumento fijo A. Vemos así que, como arriba adelanté, el signo $\smile$ funciona como signo de cuantificación universal, correspondiente a nuestro ∀, y liga la variable individual o predicativa inscrita en él. La expresión ‘*a* = *b*’ denota **v** si ‘*a*’ y ‘*b*’ denotan el mismo objeto y denota **f** si ‘*a*’ y ‘*b*’ no denotan el mismo objeto. Así, si ‘*a*’ y ‘*b*’ son términos, ‘*a* = *b*’ dice que *a* es idéntico a *b* y si ‘*a*’ y ‘*b*’ son aseveraciones, ‘*a* = *b*’ dice que *a* si y sólo si *b*. Por último, el signo ┬└ designa la función binaria definida así: (i) ┬A └B es **f** si y sólo si B es **v** y A es un objeto distinto de **v**; (ii) si A es **v** o si B no es **v**, ┬A └B es **v** (Frege 1891, p. 28; 1893, p. 20).[17]

---

de ellas estuviera definida en todo el universo de objetos. Esto es una ilusión, pues si $\mathcal{W}$ designa el universo de objetos, una función irrestricta binaria (una “función de dos variables”) no está definida en $\mathcal{W}$ sino en $\mathcal{W}^2$, una función ternaria en $\mathcal{W}^3$, etc.

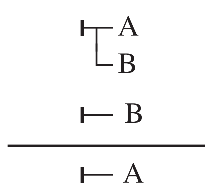
16 Esta función, como todas las de Frege, está definida para cada objeto (nota 15). Según esto, —┬—*x* es **v** aunque ‘*x*’ no denote un valor veritativo.

17 Frege 1879, p. 5, dice que “si A y B significan contenidos aseverables, hay las cuatro posibilidades siguientes: (1) se afirma A y se afirma B; (2) se afirma A y se niega B; (3) se niega A y se afirma B; (4) se niega A y se niega B. ├┬A └B significa entonces la aseveración de *que la tercera de estas posibilidades no se cumple, sino una de las otras tres*.” A la luz de esta caracterización, Baker y Hacker (1984), concluyen que en 1879 Frege no había dado aún con la definición verifuncional de la implicación, descubierta por Filón de Megara en el siglo III a.C. El texto citado respalda este diagnóstico sólo si se toma *al pie de la letra*, esto es, si se entiende que ├┬A └B es una aseveración *biográfica* relativa a lo que alguna persona indeterminada de hecho asevera o no asevera. Pero no creo que fuera la intención de Frege entenderlo así. En las pp. 10–11 del mismo libro trata como equivalentes las expresiones ‘se afirma B’ (‘B wird bejaht’) y ‘B ha de afirmarse’ (‘B zu bejahen . . . ist’), y creo que ya en 1879 Frege pensaba que un contenido aseverable *ha de* afirmarse si y sólo si es verdadero.

La última definición indica sin lugar a dudas que ─┬─A └─B representa la implicación material que hoy representamos con (A $\rightarrow$ B). Frege advierte oportunamente al lector que ─┬─A └─B corresponde a 'si B, entonces A' ("wenn A, so B") sólo *parcialmente*, en *algunos* casos.[18] Hay, en cambio, funciones compuestas de ─┬─ y ─┬─└─ que corresponden a ciertas conjunciones familiares. Así, ┬─┬─┬ └─ corresponde exactamente a la conjunción 'y', pues, como el lector puede fácilmente calcular, ┬─┬─┬─A └──B denota **v** si —A *y* —B denotan **v**, y de otro modo denota **f**. Y ──┬── └─┬ corresponde a la conjunción 'o' entendida en un sentido no excluyente, puesto que ──┬──A └─┬B denota **f** si —A *y* —B denotan **f**, y de otro modo denota **v**. En otras palabras, ┬─┬─┬ └─ es la función que simbolizamos '$\wedge$' y ──┬── └─┬ es la función que simbolizamos '$\vee$'. Aprovechando la correspondencia entre la función $\wedge$ y la conjunción 'y', los profesores de lógica suelen introducir primero esa función en sus explicaciones y luego *definen* $\rightarrow$ como una función compuesta de $\wedge$ y $\neg$. Frege no obró así porque, mucho más que traducir la escritura conceptual BS al lenguaje cotidiano, le interesaba escribir deducciones de manera perspicua y $\rightarrow$ se presta muchísimo mejor que $\wedge$ para este propósito.

En efecto, con los signos de BS y las convenciones semánticas de Frege, la regla *modus ponens* (MP) puede enunciarse así: *De las premisas* ├┬─A └─B y ├── B, *deducir la conclusión* ├── A. O, más gráficamente:

18 Su buen sentido contrasta con la irreflexión de esos profesores de lógica que enseñan que (A $\rightarrow$ B) se lee 'si A, entonces B' y tienen luego que enfrentar la mirada escéptica de sus alumnos cuando, a resultas de ello, se ven forzados a sostener que un enunciado como 'si la luna está hecha de queso, entonces 2 + 2 = 4' expresa una verdad necesaria.



Presentada así, no cabe la menor duda de que la deducción es correcta, cualesquiera que sean las oraciones declarativas A y B, puesto que, si la segunda premisa es verdadera, la primera premisa sólo puede ser verdadera si la conclusión también lo es. Para controlar la corrección de una inferencia en que se usa esta regla basta comprobar por simple inspección visual que la expresión a la derecha del signo $\vdash$ en la conclusión es igual a la expresión a la derecha de la línea horizontal superior del signo $\vdash\!\!\sqsubset$ en la primera premisa y que la expresión a la derecha de la línea horizontal inferior de este último signo es igual a la expresión a la derecha del signo $\vdash$ en la segunda premisa.

En el § 11 de *Escritura conceptual* Frege enuncia la regla de inferencia que hoy suele llamarse *generalización universal* (GU): *Deducir de la premisa* $\vdash\!\!\sqsubset$ Φ(*a*) / A *la conclusión* $\vdash\!\!\sqsubset$ 𝔞 Φ(𝔞) / A *siempre que* A *sea una expresión en la cual no figura a y que a sólo figure en* Φ(*a*) *en las posiciones que ocupa* 𝔞 *en* Φ(𝔞). Frege justifica esta regla así: "Si —𝔞— Φ(𝔞) es negado se tiene que poder asignar a *a* un significado tal que Φ(*a*) sea negado. Por lo tanto, si —𝔞— Φ(𝔞) fuera negado y A fuera afirmado, se tendría que poder asignar a *a* un significado tal que A fuese afirmado y Φ(*a*) fuese negado. Pero ello es imposible en virtud de $\vdash\!\!\sqsubset$ Φ(*a*) / A, ya que esto significa que está excluido el caso en que Φ(*a*) es negado y A es afirmado. Por lo tanto, no es posible negar —𝔞— Φ(𝔞) y afirmar A; esto es: $\vdash\!\!\sqsubset$ 𝔞 Φ(𝔞) / A" (Frege 1879, pp. 21s.)

Sin embargo, en las deducciones de ese libro, Frege no utiliza —por lo que veo— esta regla GU, sino únicamente la regla MP y una regla de sustitución, cuyos requisitos no llega a formular expresamente. Conforme a esta

última, de una dada aseveración $\mathscr{A}$ se infiere la aseveración $\mathscr{A}'$ obtenida al sustituir uniformemente una determinada expresión $\mathscr{B}$ que figura una o más veces en $\mathscr{A}$ por otra expresión $\mathscr{C}$. (Digo que $\mathscr{C}$ sustituye *uniformemente* a $\mathscr{B}$ en $\mathscr{A}$ si se escribe $\mathscr{C}$ en lugar de $\mathscr{B}$ en todas las posiciones que $\mathscr{B}$ ocupa en $\mathscr{A}$). Evidentemente, la aplicación de este procedimiento se puede controlar por simple inspección visual. Su justificación es también muy obvia cuando se manda reemplazar —como en las primeras deducciones del libro (Frege 1879, §§ 15–19)— una cursiva minúscula utilizada como variable proposicional en cierta aseveración $\mathscr{A}$ por una oración (esto es, una expresión apta para denotar cualquiera de los valores veritativos **v** y **f**): bajo las reglas semánticas de BS, tal sustitución no puede afectar el valor veritativo denotado por $\mathscr{A}$. Pero no todas las sustituciones prescritas en esa obra siguen este sencillo patrón.[19] En 1893, Frege describe y justifica en detalle los procedimientos de sustitución admisibles y enuncia además otras reglas que validan pasos que en 1879 presentaba como sustituciones. El "Resumen de reglas" en Frege 1893, § 48, enumera doce (más seis "estipulaciones sobre el uso de paréntesis"). La regla MP es sólo una de ellas. No puedo examinar aquí las otras. Baste señalar que —como vimos en el caso de MP— aunque la justificación de cada regla es semántica, su aplicación y el control de su aplicación depende exclusivamente de la apariencia visual de las expresiones envueltas, y no de su significado. Esto es esencial para el propósito de BS, que es poner de manifiesto, sin lugar a dudas, los nexos deductivos entre las verdades aritméticas y las leyes lógicas suplementadas con definiciones.

Una expresión $\mathscr{A}$ del BS representa una ley lógica si denota la verdad *en todo caso*, sea cual sea la denotación de sus partes. En la segunda parte de *Escritura conceptual*, Frege enuncia nueve expresiones que cumplen este requisito, según hace evidente mediante consideraciones semánticas.[20] Estas expresiones son las premisas indemostradas de que parten sus deducciones. Constituyen, pues, una lista de axiomas para la lógica. Las reproduzco en el

[19] En la nota 6 mencioné las sustituciones que Frege aventura para probar la Proposición 90. Hay otros ejemplos.

[20] Por ejemplo, el axioma 1 excluye el caso en que *p* es **f**, *q* es **v** y *p* es **v**, el cual es evidentemente imposible, puesto que *p* no puede ser **v** y **f** a la vez (Frege 1879, § 14). Por cierto, esto no es una *demostración* del axioma; pero al *aclarar* lo que éste significa se pone de manifiesto que no es posible negarlo.

Apéndice X. Con el simbolismo del cálculo predicativo se las puede formular como sigue:[21]

1. $(p \to (q \to p))$
2. $((p \to (q \to r)) \to ((p \to q) \to (p \to r)))$
8. $((p \to (q \to r)) \to (q \to (p \to r)))$
28. $((p \to q) \to (\neg q \to \neg p))$
31. $(\neg\neg p \to p)$
41. $(p \to \neg\neg p)$
52. $((c = d) \to (\Phi(c) \to \Phi(d)))$
54. $(c = c)$
58. $(\forall x \Phi(x) \to \Phi(a))$

La fuerza de estos axiomas depende, por cierto, de las reglas de inferencia que se adopten. Si las letras *p*, *q* y *r* se tratan como variables proposicionales, libremente sustituibles por oraciones, los Axiomas 1, 2, 8, 28, 31 y 41, combinados con la regla MP, bastan para deducir todas las oraciones compuestas que denotan la verdad cualquiera que sea el valor veritativo de las oraciones componentes (esto es, todas las *tautologías*, en el sentido del Apéndice IX.C). Supongamos, además, que en el Axioma 58 se permite la sustitución de *a* por cualquier expresión $\tau$ que designe un objeto y de $\Phi(\ )$ por una expresión insaturada que se convierte en una oración en cuanto sus "blancos" se llenen todos con $\tau$. Entonces, como demostrará Gödel (1930), los Axiomas 1, 2, 8, 28, 31, 41 y 58, combinados con las reglas MP y GU bastan para deducir todas las verdades lógicas expresables en un cálculo predi-

[21] Doy a cada axioma el número que tiene en el libro de Frege (los números intermedios corresponden a teoremas deducidos de los axiomas precedentes). La "traducción" ofrecida demanda algunos comentarios. Debemos recordar que en BS, el signo '=' ('≡' en 1879) funciona como nuestro símbolo de identidad '=' si las expresiones a la izquierda y la derecha son variables individuales o nombres de objetos, y como nuestro símbolo de equivalencia '↔' si dichas expresiones son oraciones. Frege permite sustituir las letras *c* y *d* que figuran en los Axiomas 52 y 54 tanto por nombres como por oraciones. Por último, como señalé en la nota 6, en 1879 Frege se permite reemplazar la variable ligada en el Axioma 58 no sólo por nombres, sino también por *predicados*.

cativo de primer orden sin identidad (como el presentado en el Apéndice IX.A, B). Para deducir todas las verdades lógicas expresables en el cálculo predicativo de primer orden *con identidad* hay que agregar los Axiomas 52 y 54 —en el entendido de que $\Phi(d)$ en el 52 se obtiene por sustitución uniforme *o parcial* de $c$ por $d$ en $\Phi(c)$— y autorizar la libre sustitución de las letras $c$ y $d$ que figuran en ellos por expresiones que designen objetos. Pero, como pondrá en claro el mismo Gödel (1930a, 1931), este sistema de axiomas y reglas de inferencia, suplementado con todas las definiciones que se deseen, no es suficiente para deducir todas las verdades aritméticas. Distinto es el caso si adoptamos MP y la loca regla de sustitución empleada por Frege en 1879, o los axiomas y reglas de inferencia postulados por Frege en 1893. Tales recursos sí que bastan para deducir las verdades aritméticas y —en general— todo lo que uno quiera, pues, como Russell le informó a Frege el 16 de junio de 1902, el sistema es inconsistente.[22]

[22] En una carta dirigida a Frege en esa fecha, Russell vincula directamente la paradoja que hoy se conoce por su nombre a la liberalidad con que Frege admite y maneja las variables predicativas en *Escritura conceptual*. Esa liberalidad se manifiesta específicamente en la práctica fregeana —a que aludí en la nota 6— de sustituir por un predicado la variable $c$ que figura en la línea superior del Axioma 58, . Sea $w$ el predicado "…es un predicado no predicable de sí mismo". Russell se pregunta si $w$ es predicable de sí mismo. Como quiera que se conteste esta pregunta, se obtiene una contradicción. Por lo tanto, concluye Russell, $w$ no es un predicado. Pero la contradicción es inevitable bajo las reglas (o prácticas) de *Escritura conceptual*. En efecto, en el BS de 1879, la condición con que Russell caracteriza el predicado $w$ tendría que escribirse así: $\vdash\!\overset{\mathfrak{F}}{\smile}\!-\ ((\text{—}\ w(\mathfrak{F})) \equiv (-\!\!\top\!\!-\ \mathfrak{F}(\mathfrak{F})))$. Si en el Axioma 58 sustituimos, a la manera de Frege, $f(\Gamma)$ por $((\text{—}\,w(\Gamma)) \equiv (-\!\!\top\!\!-\,\Gamma(\Gamma)))$, $\mathfrak{a}$ por $\mathfrak{F}$ y $c$ por $w$, y aplicamos MP al par de premisas que forman el resultado de esta sustitución y la condición antedicha obtenemos inmediatamente la contradicción

$$\vdash ((\text{—}\ w(w)) \equiv (-\!\!\top\!\!-\ w(w)))$$

Russell agrega: "Asímismo, no hay ninguna clase (como un todo) de aquellas clases que —como todos— no se pertenecen a sí mismas. De ello concluyo que bajo ciertas circunstancias un conjunto definible no forma un todo (*eine definierbare Menge kein Ganzes bildet*)." (Frege, WB, p. 211). La paradoja de Russell surge, por cierto, también en el sistema más riguroso de Frege 1893/1903 (véase el Apéndice XII).

## 2.3 FUNDAMENTOS DE LA ARITMÉTICA

El aporte más importante de Frege a la filosofía matemática consiste en lo que ya hemos visto: su concepción de la escritura conceptual y su análisis de las relaciones y de las aseveraciones condicionales y universales (Capitulo 2.2). Su fundamentación de la aritmética con estos medios impresiona todavía a ciertos filósofos, pero los matemáticos le han prestado muy poca atención. En todo caso, aunque no nos detengamos a examinar el formalismo de sus deducciones, vale la pena darle una mirada a las definiciones que introdujo para "reducir" los conceptos específicos de la aritmética a ideas básicas comunes a todas las áreas del pensamiento. La motivación y alcance de la empresa de Frege —y también la peculiar ceguera que la vicia— se apreciarán mejor si examinamos primero la obra contemporánea —mucho más influyente y, en mi opinión, más acertada— de Richard Dedekind y Giuseppe Peano.[1]

### *2.3.1 Peano (1889)*

En 1889 Peano publicó la monografía *Arithmetices principia, novo methodo exposita*. En el largo prefacio presenta y explica la escritura conceptual inventada por él, que es el antecedente histórico directo de la que usamos hoy. El § 1 contiene la primera versión publicada de los célebres Axiomas de Peano, que constituyen justamente el sistema de axiomas para la aritmética elemental cuya consistencia Hilbert y sus colaboradores intentarán demostrar. Peano agradece a Boole, Schröder, Jevons, Peirce y MacColl sus aportes a la lógica y cita también el libro sobre aritmética de Dedekind (1888) que comentaré enseguida, pero no nombra a Frege, lo que me hace suponer

1 Al redactar este capitulo tuve a la vista con mucho provecho la obra de Gillies (1982) y la Sección XIX del libro de Crispin Wright (1983).

que en 1889 todavía no lo conocía.[2] Reproduzco a continuación los Axiomas de Peano, en la notación original (doy entre paréntesis una versión castellana):[3]

**P1** $1 \; \varepsilon \; \mathrm{N}.$
(1 es un número.)

**P2** $a \; \varepsilon \; \mathrm{N}. \supset . \; a + 1 \; \varepsilon \; \mathrm{N}.$
(Si *a* es un número, el siguiente de *a* es un número.)

**P3** $a, b \; \varepsilon \; \mathrm{N}. \supset . \; a = b. = . \; a + 1 = b + 1.$
(Si *a* y *b* son números, *a* es igual a *b* si y sólo si el siguiente de *a* es igual al siguiente de *b*.)

**P4** $a \; \varepsilon \; \mathrm{N}. \supset . \; a + 1 \; -\!= 1.$
(Si *a* es un número, el siguiente de *a* no es igual a 1.)

**P5** $k \; \varepsilon \; \mathrm{K}. \supset \therefore 1 \; \varepsilon \; k \therefore x \; \varepsilon \; \mathrm{N}. \; x \; \varepsilon \; k : \supset_x . \; x + 1 \; \varepsilon \; k :: \supset . \; \mathrm{N} \supset k.$
(Si *k* es una clase tal que (i) 1 pertenece a *k* y (ii) para todo objeto *x*, si *x* es un número y *x* pertenece a *k*, también el siguiente de *x* pertenece a *k*, entonces *k* incluye a toda la clase de los números.)

2 En un artículo sobre "Los principios de la lógica matemática" aparecido dos años más tarde, Peano cita a Frege 1879, pero sólo para informar al lector que, "en vez de $a \supset b$, [Frege] escribe $\vdash\!\!\!\begin{array}{l} a \\ b \end{array}$" (Peano 1891, n. 5). El 30 de enero de 1894, contestando a una carta de Frege, Peano le comunica que "j'ai acheté il y a quelques temps vos *Die Grundlagen der Arithmetik* [Frege 1884]" y el 3 de octubre de 1896 le cuenta que acaba de releer *Begriffschrift* (1879) y *Grundgesetze* (tomo I, 1893), "avec nouveau plaisir" (Frege, WB, p. 177, 189), pero no sabemos cuando leyó estas obras por primera vez. Peano publicó en 1895 una reseña del tomo I de *Grundgesetze* que dio lugar a una interesante respuesta de Frege.

3 En el libro de Peano los axiomas P2–P5 llevan los números 6, 7, 8 y 9, respectivamente. Los números 2–5 corresponden a axiomas que gobiernan el uso *entre números* del signo '=', el cual se emplea a la vez como símbolo de la equivalencia entre aseveraciones y de la igualdad aritmética. Peano (1898) enuncia sólo cinco axiomas de la aritmética, correspondientes a **P1**–**P5** (supongo que esta simplificación se debe a que llegó a pensar como Frege que los axiomas que gobiernan el símbolo '=' en ambos usos pertenecen a la lógica general; cf. nota 5). Allí emplea el símbolo '0' en vez de '1' para designar al miembro distinguido de la clase N, esto es, el objeto que según **P1** *es un número* y según **P4** *no es igual al siguiente* de ningún número; con ello se evita la perniciosa confusión entre '1' y '+1' que comento en la nota 4.

En los enunciados de Peano figuran algunos signos de uso general que ha explicado en la presentación de su escritura conceptual, a saber, $-$, $=$, $\varepsilon$, Ɔ, $Ɔ_x$, la constante K y los puntos. Hay además tres símbolos nuevos, no definidos, que son los primitivos de su aritmética, a saber, N, 1 y la expresión +1.[4] Antes de examinar la índole sintáctica de estos primitivos y las condiciones semánticas que les imponen los axiomas, conviene hacer algunas observaciones sobre los signos de uso general. Peano dice que "K significa *clase* o agregado de entes". Los puntos funcionan como los paréntesis en el álgebra, entendiéndose que, en una fórmula dividida por puntos, se asocian, ante todo, los signos que no están separados por ningún punto; enseguida, los que están separados por un punto; luego, los que están separados por dos puntos, etc. Así, $ab.cd:ef.gh\therefore k$ significa lo mismo que $(((ab)(cd))((ef)(gh)))k$. Aunque Peano no lo dice, debemos entender que la simple yuxtaposición de dos oraciones expresa la conjunción de éstas.

Los demás signos de uso general corresponden, aunque no sin equívocos, a otros que utilizamos hoy. El signo =, escrito entre nombres de objetos, expresa la identidad de los objetos nombrados, como nuestro signo de la misma forma; pero Peano también lo escribe entre oraciones para expresar la equivalencia de éstas, como nuestro signo ↔ (cf. la segunda posición que

4 También el signo '=' es un primitivo de la aritmética en Peano 1889 (además de servir como símbolo lógico de equivalencia), pero para nuestros propósitos es preferible no hacer caso de esta peculiaridad (por eso he omitido los axiomas "aritméticos" 2–4 que atribuyen a la igualdad *entre números* la consabida reflexividad, simetría y transitividad, y el axioma 5, que estipula que si $a = b$ y $a$ es un número, entonces $b$ es un número). Es importante advertir que, aunque consta de dos caracteres, '+1' es un solo símbolo simple que representa la función unaria 'el siguiente de…'. El signo '1' que vemos en él es sólo *tipográficamente* igual a la constante individual '1' que se usa en el axioma **P1**, y el signo '+' que lo precede *no representa* aquí la adición, es decir, la función binaria que asigna a cada par de números $\langle x,y\rangle$ el número que es la suma de $x$ e $y$. Desafortunadamente, Peano utiliza en su libro el mismo signo '+' como símbolo de la adición, que define así en la Proposición 18:

$$a, b \text{ e N.} Ɔ . a + (b + 1) = (a + b) + 1$$

Esto es: Si $a$ y $b$ son números, la suma de $a$ y el siguiente de $b$ es igual al siguiente de la suma de $a$ y $b$. Nótese que en esta definición el primer y el tercer signo '+' significan *la suma* (de los números que figuran a cada lado), pero el segundo y el cuarto se funden con el signo '1' para significar *el siguiente* (del último número a su izquierda).

el signo = ocupa en **P3**).[5] El signo **−** significa negación, como nuestro signo ¬; pero también expresa la operación booleana de formar el complemento (si *a* es una clase, **−***a* es la clase formada por todos los objetos que no pertenecen a *a*). La sintaxis del signo de negación difiere de la actual: si *a* y *b* son objetos, '*a* **−**= *b*' significa '*a* no es idéntico a *b*', lo que ahora escribiríamos '¬(*a* = *b*)' o, más informalmente, '$a \neq b$'.

El signo ε es, por cierto, el antecedente histórico directo de nuestro signo ∈: en **P5** la cláusula '1 ε *k*' dice que el objeto llamado 1 pertenece a —es uno de los elementos de— la clase o colección de objetos llamada *k*. Pero Peano explica que ε debe leerse 'es' (ἐστί en griego; *è* en italiano). Cabría, pues, entender este signo meramente como una cópula que une el sujeto mencionado a la izquierda con el predicado escrito a la derecha. Bajo esta interpretación la cláusula '*k* ε K' en **P5** puede leerse '*k* es clase', sin que haga falta suponer que hay una *clase de todas las clases* denotada por la constante K.

Peano fue el primero en distinguir entre la relación de pertenencia a una clase (o conjunto), simbolizada por ε, y la relación de inclusión entre clases (o conjuntos), simbolizada por Ɔ. Si *a* y *b* son clases, entonces *a* Ɔ *b* si y sólo si, cualquiera que sea el objeto *x*, *x* ε *a* implica que *x* ε *b*. En esta función, Ɔ corresponde pues a nuestro signo ⊆. Pero Ɔ desempeña además otra función: si *a* y *b* son oraciones, entonces '*a* Ɔ *b*' significa, según Peano, que *b* se deduce de *a*. Esta explicación no es muy iluminadora ya que, debido a la ausencia total de reglas de inferencia explícitas, no es fácil determinar cuándo una oración se deduce de otra en su sistema. Con todo, Peano indica —en un pasaje citado en la nota 2— que si *a* y *b* son oraciones, '*a* Ɔ *b*' significa en verdad lo mismo que ahora expresamos con '$a \to b$' (el lector percibirá la semejanza entre el signo Ɔ y la "herradura" ⊃ utilizada por Russell y otros en vez de la flecha →). El doble sentido del signo Ɔ es notorio en **P5**, pero también en la definición misma de Ɔ como signo de inclusión (Peano 1889, Prefacio, Proposición 50):

[5] Como vimos en la p. 142, n. 21, también Frege utiliza su signo de identidad = para expresar la equivalencia entre oraciones. Pero en su sistema esto no constituye un uso equívoco de dicho signo, ya que, según él, una oración denota su valor veritativo. Obviamente, si dos oraciones son equivalentes, el valor veritativo que ambas denotan es uno y el mismo.

$$a, b \;\varepsilon\; \mathrm{K}. \supset \therefore a \supset b : = : x \;\varepsilon\; a. \supset_x. x \;\varepsilon\; b \qquad (50)$$

Aquí $\supset$ significa inclusión cuando está escrito entre los nombres de clase *a* y *b*, pero significa implicación en los otros casos. Con el subíndice *x* que acompaña a la última implicación Peano quiere decir que ésta vale cualquiera que sea el objeto *x*. El subíndice *x* hace, pues, las veces de nuestro cuantificador universal $\forall x$ (en lugar de '$x \;\varepsilon\; a. \supset_x. x \;\varepsilon\; b$' hoy escribiríamos '$\forall x(x \in a \to x \in b)$). Peano escribe asimismo $=_x$ para generalizar una equivalencia.

Las explicaciones precedentes, aunque tediosas, apuntan a una importante limitación común a toda escritura conceptual. Aunque se pretenda absoluta, la precisión que tal escritura confiere al pensamiento no será mayor que la que se logre discurriendo —para definirla y reglamentarla— en una lengua viva. Por otra parte, la escritura conceptual permite fijar ideas de *un cierto grado* de precisión y protegerlas —merced a la artificialidad y carencia de poder sugestivo de sus símbolos— contra la erosión por el torrente de paráfrasis y metáforas de que se nutre el discurso "natural".

Atendamos ahora a los primitivos N, 1 y +1. En los axiomas **P1**–**P4**, N figura siempre a la derecha del signo ε. Por lo tanto, según la interpretación arriba mencionada, podríamos entender que N es un predicado, no el nombre de una clase o colección existente de objetos. Pero en **P5** figura la cláusula $\mathrm{N} \supset k$. De acuerdo con la Proposición 50 del Prefacio, esto sólo quiere decir algo si N ε K. El axioma **P1** dice que 1 es un elemento distinguido de esta clase N. Los axiomas **P2**–**P4** caracterizan la aplicación de N en N por $x \mapsto x + 1$. **P2** indica que se trata efectivamente de una aplicación con dominio y codominio N, puesto que dice que para cada objeto *a* que pertenezca a N hay también en N un determinado elemento designado por $a + 1$. **P3** indica que esta aplicación es inyectiva ($a + 1 = b + 1$ sólo si $a = b$).[6] **P4** dice que el objeto distinguido 1 no cae dentro del alcance de la aplicación $x \mapsto x + 1$ (no existe un $x \;\varepsilon\; \mathrm{N}$ tal que $1 = x + 1$). Finalmente, **P5** dice que la

6 Según lo que llevamos dicho, la apódosis $a = b. = . a + 1 = b + 1$ de **P3** equivale a la conjunción de las dos condiciones $a + 1 = b + 1. \supset . a = b$ y $a = b. \supset . a + 1 = b + 1$. La primera de estas condiciones establece que la aplicación $x \mapsto x + 1$ es inyectiva. La segunda condición es superflua si entendemos, como parece natural, que la expresión $a + 1$ es un término y, como tal, designa un y sólo objeto.

clase N está incluida en cualquier clase *k* que reúna los dos requisitos siguientes: (i) *k* contiene el elemento distinguido 1 y (ii) si *x* es cualquier objeto contenido a la vez en N y en *k*, *k* contiene también el objeto $x + 1$ (que ciertamentemente existe, en virtud de **P2**, puesto que $x \,\varepsilon\, \mathrm{N}$). Por lo tanto, si entendemos —como parece haber entendio Peano— que cada propiedad atribuible a objetos determina una clase formada por los objetos que tienen esa propiedad, el axioma **P5** justifica un método para demostrar que una propiedad es común a los objetos de la clase N. Sea *k* la clase formada por los objetos que tienen la propiedad *k**. En virtud de **P5**, para probar que todos los objetos de la clase N tienen la propiedad *k** (de modo que $\mathrm{N} \supset k$) basta probar que 1 la tiene y que, si la tiene cualquier $x \,\varepsilon\, \mathrm{N}$ también la tiene el respectivo objeto $x + 1$. Este método de prueba se llama inducción matemática (finita).

Para apreciar mejor el potencial de los axiomas de Peano, consideremos un conjunto no vacío cualquiera $\mathcal{N}$ y una aplicación inyectiva $\sigma\colon \mathcal{N} \to \mathcal{N}$, tal que un determinado objeto $\mathbf{1} \in \mathcal{N}$ queda fuera del alcance de $\sigma$ ($\forall x(x \in \mathcal{N} \to \sigma(x) \neq \mathbf{1})$). La estructura $\langle \mathcal{N},\sigma,\mathbf{1} \rangle$ evidentemente satisface los axiomas **P1**–**P4**, si entendemos que $\mathcal{N}$ es la clase N, que $\mathbf{1}$ es el objeto distinguido 1 y que $\sigma$ es la aplicación $x \mapsto x + 1$. ¿Qué condiciones adicionales tiene que cumplir $\langle \mathcal{N},\sigma,\mathbf{1} \rangle$ para que satisfaga además el axioma **P5**? En *Was sind und sollen die Zahlen?* (1888), Dedekind había dado a esta pregunta una respuesta que Peano no incorpora a su tratamiento del asunto, posiblemente porque cuando redactó *Arithmetices Principia* (1889) todavía no había estudiado bien aquel libro (que cita, sin embargo, en su Prefacio). Pero antes de hablar de ella, conviene tener presentes algunos ejemplos de estructuras que efectivamente satisfacen los cinco axiomas **P1**–**P5**. Sea $\mathcal{N}$ el conjunto de los símbolos con que representamos los enteros positivos: $\mathcal{N} = \{1, 2, 3, \ldots\}$; llamemos $\mathbf{1}$ al símbolo 1 y $\sigma$ a la aplicación que asigna a cada símbolo el símbolo siguiente.[7] Entonces $\langle \mathcal{N},\sigma,\mathbf{1} \rangle$ satisface los cinco axiomas de Peano. Pero

7 Nótese que digo *los símbolos*, no los objetos (números) que ellos supuestamente representan. Qué sean estos objetos es lo que Peano y Dedekind buscaban establecer. Para que el conjunto de símbolos $\{1, 2, 3, \ldots\}$ esté bien definido es preciso, claro está, disponer de una regla que determine la figura del sucesor inmediato de cualquier símbolo dado. Pero evidentemente disponemos de ella. ¿No sabe acaso el lector escribir en el acto el número que sigue a 37.045.622.876.359? Con un pequeño esfuerzo puede también sin duda nombrarlo, en castellano, en inglés y en cuántas lenguas conozca.

$\langle \mathcal{N},\sigma,\mathbf{1}\rangle$ también satisface los cinco axiomas de Peano si $\mathcal{N} = \{0, 1, 2, \ldots\}$ y σ es como antes, pero **1** es el símbolo 0. Fijémosnos por el momento en esta última interpretación. Sea $\mathcal{N}_n \subseteq \mathcal{N}$ el conjunto $\{0, n, 2n, 3n, \ldots\}$, **1** el 0 y $\sigma_n$ la aplicación $\mathcal{N}_n \rightarrow \mathcal{N}_n$ definida por $\sigma_n(x) = x + n$. Basta una breve reflexión para convencerse de que los axiomas **P1**–**P5** están realizados en la estructura $\langle \mathcal{N}_n,\sigma_n,\mathbf{1}\rangle$, si entendemos que la clase N de Peano es $\mathcal{N}_n$, que el objeto distinguido que él llama 1 es nuestro **1** (esto es, el símbolo 0), y que su aplicación inyectiva $x \mapsto x + 1$ es $\sigma_n$.

### *2.3.2 Dedekind (1888)*

En la obra citada, Dedekind quiso mostrar que "el concepto de número es enteramente independiente de las representaciones o intuiciones del espacio y del tiempo" y que puede considerárselo "un producto (*Ausfluß*) inmediato de las puras leyes del pensamiento" (1893, p. vii).[8] Estas palabras recuerdan el programa de Frege: probar que las verdades aritméticas se deducen de definiciones y leyes de la lógica. En el prólogo a la segunda edición Dedekind dice que leyó el libro de Frege, *Grundlagen der Arithmetik* (1884), un año

---

Podríamos, pues, identificar $\mathcal{N}$ con cualquiera de los siguientes conjuntos de *palabras*: {uno, dos, tres, . . . }, o {one, two, three, . . . }, o {eins, zwei, drei, . . . }, etc., y obtener otras tantas realizaciones de los axiomas de Peano. A primera vista, no parece que pueda sacarse mucho provecho del axioma **P5** para demostrar propiedades del conjunto $\mathcal{N}$ así interpretado, pues obviamente no hay ninguna conexión necesaria entre la figura gráfica o fónica de uno de los símbolos o palabras en cuestión y la de su sucesor inmediato. Pero las propiedades interesantes que se prueban mediante **P5** involucran funciones como la suma y la multiplicación, cuyas definiciones fundan tales conexiones. Se advertirá de paso que en cualquiera de estas interpretaciones $\mathcal{N}$ no es un conjunto de objetos físicos —paquetes de ondas acústicas, inscripciones hechas con tiza o con lápiz— sino ideales: los prototipos pensables de tales símbolos o palabras. De hecho, la inmensa mayoría de esos prototipos no ha tenido nunca ni podría tener una realización física.

8 No he tenido en mis manos un ejemplar de la primera edición de Dedekind 1888, sino sólo de la segunda (inalterada) de 1893. Por eso —y para facilitar la consulta de otras ediciones— mis citas remiten a los números (del 1 al 172) en que está dividido el texto en todas ellas. Esta convención, empero, no es aplicable al prólogo.

después de publicada la primera del suyo (1888), y que, considerando “la precisión con que el autor se expresa sobre la inferencia de $n$ a $n + 1$”, es claro que “aquí pisamos sobre el mismo terreno” (Dedekind 1893, p. xvii). Sin embargo, como veremos, hay una profunda diferencia entre ambos. Mucho más cerca está Dedekind de Cantor, con quien, como sabemos, tenía correspondencia (Capítulo 1.4). Dedekind parte introduciendo la noción de sistema (*System*), que yo no sabría distinguir de lo que Cantor llamaba ‘conjunto’ (*Menge*).[9] En aras de “la uniformidad de la dicción”, admite la existencia de sistemas “que constan de un objeto único” (esto es, de conjuntos unitarios); pero descarta explícitamente “el sistema vacío, que no contiene ningún elemento [ . . . ], aunque para otras investigaciones puede ser cómodo inventarlo” (Dedekind 1888, #2).

Básico es asímismo el concepto de *aplicación* (*Abbildung*), que Dedekind introduce así:

> Por *aplicación* φ de un sistema *S* se entiende una ley en virtud de la cual a cada elemento determinado *s* de *S pertenece* (*gehört*) una cosa determinada que se llama la *imagen* (*Bild*) de *S* y se designa con φ(*s*). Decimos también que φ(*s*) *corresponde* (*entspricht*) al elemento *s*, que φ(*s*) *surge* de *s* o *es generado* desde *s* por la aplicación φ, y que por ella *s se transforma en* φ(*s*) (*übergeht in* φ(*s*)). Si *T* es una parte cualquiera de *S*, la aplicación φ de *S* contiene a la vez una determinada aplicación de *T* que, en aras de la simplicidad, podemos designar con el mismo signo φ y que consiste en que a cada elemento *t* del sistema *T* le corresponde la misma imagen φ(*t*) que *t* posee como elemento de *S*. También el sistema que consta de todas las imágenes φ(*t*) se llamará la *imagen* de *T* y se designará con φ(*T*). Esto explica de paso el significado de la expresión φ(*S*).
>
> (Dedekind 1888, #21)

[9] He aquí una traducción del pasaje pertinente:

> En lo sucesivo entiendo por *cosa* (*Ding*) a cada objeto de nuestro pensamiento. [ . . . ] Ocurre muy a menudo que diversas cosas *a*, *b*, *c*, . . . , consideradas por cualquier motivo bajo un punto de vista común, son reunidas en la mente (*im Geiste zusammengestellt werden*) y se dice entonces que forman un *sistema S*; las cosas *a*, *b*, *c*, . . . son los *elementos* del sistema *S*, están *contenidas* en *S*. Por su parte, *S consta* de (*besteht aus*) estos elementos. Como objeto de nuestro pensamiento, un tal sistema *S* [ . . . ] es también una cosa; está completamente determinado cuando respecto de cada cosa está determinado si ella es o no es un elemento de *S*.
>
> (Dedekind 1888, ##1,2)

Este concepto de Dedekind concuerda no sólo en el nombre con lo que en alemán ahora se llama *Abbildung* (‘aplicación’† en castellano). Sorprende un tanto el aserto de que una aplicación es una *ley*, pues las aplicaciones a que Dedekind de hecho se refiere en su libro pueden ser —como las nuestras— correspondencias enteramente arbitrarias. Igual que en la matemática actual, una aplicación está individualizada aquí por su *dominio*, esto es, por el particular sistema a cuyos elementos ella asigna “imágenes” (en contraste con las funciones de Frege que son, por así decir, aplicaciones del universo entero). Por eso Dedekind distingue con primor inusitado en su tiempo entre una aplicación φ definida en un conjunto $S$ y la restricción de φ al subconjunto $T \subset S$ (aunque, “en aras de la simplicidad”, usa la misma letra φ para designar a ambas). El lector echará de menos en el pasaje trascrito una referencia al sistema de donde la aplicación φ toma sus imágenes, o sea lo que hoy llamamos el *codominio* de φ. Dedekind introduce este concepto más adelante y enseña a describir una aplicación —como se hace hoy—nombrando su dominio *y* su codominio:

> Si φ es una [ . . . ] aplicación de un sistema $S$ y $\varphi(S)$ es una parte de un sistema $Z$ llamamos a φ una aplicación de $S$ *en* $Z$, y decimos que $S$ es aplicada por φ en $Z$.
>
> (Dedekind 1888, #36)

Pero no insiste —como haríamos hoy— en que el codominio de una aplicación también contribuye a individualizarla.[10]

Dedekind llama ‘aplicación similar’ (*ähnliche Abbildung*) a lo que hoy llamamos una aplicación *inyectiva*. Dos sistemas son similares si uno de ellos

[10] Esto se debe, seguramente, a la índole misma de las aplicaciones que Dedekind considera en su libro. El codominio $B$ de una aplicación $f\colon A \to B$ cobra importancia como factor individualizador cuando atendemos a las propiedades inherentes a $f$ en virtud de las estructuras características de $A$ y $B$. Por ejemplo, si $\mathcal{S}$ es una superficie curva en el espacio euclidiano $\mathcal{E}$, la métrica estándar de $\mathcal{E}$ induce una métrica en $\mathcal{S}$ (que determina, por ejemplo, cuales son las rutas más cortas que puede seguir una hormiga que camina sobre $\mathcal{S}$). La inclusión $\boldsymbol{\iota}\colon \mathcal{S} \to \mathcal{E}$, $x \mapsto x$, que asigna a cada punto $x$ de la superficie $\mathcal{S}$ el mismo punto considerado como elemento del espacio $\mathcal{E}$, difiere entonces de la identidad $I_{\mathcal{S}}\colon \mathcal{S} \to \mathcal{S}$, $x \mapsto x$, puesto que $I_{\mathcal{S}}$ es una isometría, pero $\boldsymbol{\iota}$ normalmente no lo es (dos pares de puntos equidistantes en $\mathcal{S}$ generalmente no equidistan en $\mathcal{E}$).

es la imagen de una aplicación similar del otro. En otras palabras, el sistema $S_1$ es *similar* al sistema $S_2$ si hay una aplicación inyectiva $\varphi$: $S_1 \rightarrow S_2$ tal que $\varphi(S_1) = S_2$. En tal caso, decimos hoy que $\varphi$ es *biyectiva*. Evidentemente, si $S_1$ es *similar* a $S_2$, ambos sistemas son equinumerosos. Valiéndose de este concepto, Dedekind construye su célebre definición de infinito: Un sistema $S$ es infinito si y sólo si es similar a una parte propia de sí mismo, esto es, si y sólo si hay una aplicación inyectiva $\varphi$: $S \rightarrow S$ tal que $\varphi(S) \neq S$.[11] Dedekind "demuestra" el siguiente "teorema": Hay sistemas infinitos (1888, #66).[12]

Para explicar cómo Dedekind concibe la aritmética tenemos que definir el término 'cadena'. Sea $S$ un sistema cualquiera y $\varphi$ una aplicación de $S$ en $S$. Una parte $K$ de $S$ es una *cadena* (*Kette*) con respecto a $\varphi$ si $K$ incluye su

11 Intuitivamente diríamos que un conjunto $S$ es infinito si y sólo si ningún número natural es suficientemente grande para enumerar a $S$, esto es, si y sólo si $|S| \geqslant |\omega|$. Pero Dedekind quiso tener un concepto de infinito utilizable en la caracterización de los números naturales. Es fácil comprobar que todo conjunto infinito en el sentido de Dedekind también es infinito en el sentido intuitivo (pues necesariamente incluye, como se verá más adelante, un conjunto equinumeroso con $\omega$). Mas para establecer la *equivalencia* de ambos conceptos se requiere el Axioma de Selección. Dedekind lo utiliza de hecho implícitamente en su demostración de que cualquier conjunto que sea equinumeroso con $\omega$ puede biyectarse sobre una parte propia suya (1888, #159).

12 Dedekind razona así: Sea $G$ el sistema de todas las cosas que pueden ser objeto de mi pensamiento. Si $g$ es un elemento cualquiera de $G$, llamaré $g'$ al pensamiento de que $g$ puede ser objeto de mi pensamiento. Obviamente, $g'$ también es un elemento de $G$. La aplicación $\varphi$: $g \mapsto g'$ es una inyección de $G$ en $G$ (ya que, si $g \neq t$, el pensamiento de que $g$ puede ser objeto de mi pensamiento es distinto del pensamiento de que $t$ puede ser objeto de mi pensamiento). No todo elemento de $G$ pertenece a $\varphi(G)$ (ya que hay objetos posibles de nuestro pensamiento que no son pensamientos acerca de otros objetos). Por lo tanto, $G$ es infinito. (Boolos 1990 hace interesantes observaciones sobre este argumento de Dedekind).

Años más tarde, Dedekind retirará esta "prueba" porque la suposición de que existe el sistema $G$ lleva directamente a las paradojas de la teoría de conjuntos, aunque expresaba todavía la confianza en que "una investigación rigurosa de la capacidad creadora del espíritu para formar con elementos determinados un nuevo [objeto] determinado, su sistema, que necesariamente se distingue de cada uno de esos elementos, conducirá sin duda a una reformulación de los fundamentos de mi escrito que supere todas las objeciones" (citado por Webb 1980, p. 63). Como vimos en el Capítulo 1.8, los conjuntistas del siglo XX simplemente *postulan* que existe al menos un conjunto infinito. Es más claro y honesto.

propia imagen $\varphi(K)$. En otras palabras, un sistema $K$ es una cadena con respecto a una aplicación $\varphi$: $S \to S$ si y sólo si $\varphi(K) \subseteq K \subseteq S$. En aras de la brevedad diré '$\varphi$-cadena' en vez de 'cadena con respecto a $\varphi$'. Obviamente, el propio sistema $S$ es una $\varphi$-cadena. Consideremos ahora una parte cualquiera $A \subseteq S$. *La $\varphi$-cadena de A* —designada en el libro de Dedekind por $A_0$— es la intersección de todas las $\varphi$-cadenas que incluyen a $A$. Obsérvese que, si $A$ mismo es una cadena, $A_0 = A$. Si $x$ es un *elemento* cualquiera de $S$, llamaré 'la $\varphi$-cadena de $x$' y designaré —como Dedekind— con $x_0$ a la $\varphi$-cadena del sistema $\{x\}$ cuyo único elemento es $x$. (Lamentablemente, la notación $A_0$, $x_0$, no indica la aplicación con respecto a la cual estos sistemas son cadenas; pero la usaré sólo en contextos en que no hay peligro de confusión).[13]

Dedekind demuestra el siguiente Teorema de la Inducción Completa (*Satz der vollständigen Induction*; en adelante, TIC): Sean $\varphi$: $S \to S$, $A \subseteq S$ y $A_0$ la $\varphi$-cadena de $A$; sea $\Sigma$ un sistema cualquiera (incluido o no en $S$). Para que $A_0$ esté incluida en $\Sigma$ ($A_0 \subseteq \Sigma$) son suficientes estas dos condiciones: (i) $A \subseteq \Sigma$ y (ii) $x \in A_0 \cap \Sigma$ sólo si $\varphi(x) \in \Sigma$. En efecto, si se cumple la condición (i), $A \subseteq A_0 \cap \Sigma$, de modo que el sistema $G = A_0 \cap \Sigma$ existe.[14] Tenemos que $A \subseteq G \subseteq A_0$. Como $A_0$ es una $\varphi$-cadena, $x \in A_0 \to \varphi(x) \in A_0$. Por lo tanto, $\varphi(G) \subseteq A_0$. Por otra parte, (ii) implica que $\varphi(G) \subseteq \Sigma$. Por lo tanto, si (ii) es verdad, tenemos que $\varphi(G) \subseteq G$. $G$ es, pues, una $\varphi$-cadena, una de las $\varphi$-cadenas que incluyen a $A$. Por lo tanto, $A_0 \subseteq G \subseteq \Sigma$. Q.E.D.

13 Anoto un resultado al que me referiré más adelante: Cualesquiera que sean la aplicación $\varphi$: $S \to S$ y el subsistema $A \subseteq S$, $\varphi(A_0) = (\varphi(A))_0$: la $\varphi$-imagen de la $\varphi$-cadena es la $\varphi$-cadena de la $\varphi$-imagen (Dedekind 1888, #57). Para probarlo, demostramos primero el siguiente lema: Si hay una $\varphi$-cadena $L$ tal que $\varphi(A) \subseteq L$, también hay una $\varphi$-cadena $K$ tal que $A \subseteq K$ y $\varphi(K) \subseteq L$. En efecto, sea $K = A \cup L$. Entonces $A \subseteq K$ y $\varphi(K) = \varphi(A) \cup \varphi(L)$. Como $\varphi(A) \subseteq L$ (por hipótesis) y $\varphi(L) \subseteq L$ (porque $L$ es una $\varphi$-cadena), $\varphi(K) \subseteq L \subseteq K$. Por lo tanto, $K$ es una $\varphi$-cadena. Sea, pues, $(\varphi(A))_0 = L$. Entonces $L$ es una $\varphi$-cadena tal que $\varphi(A) \subseteq L$ y —en virtud del lema— existe una $\varphi$-cadena $K$ tal que $A \subseteq K$ y $\varphi(K) \subseteq L$. Por lo tanto, $A_0 \subseteq K$ (ya que $A_0$ es la intersección de todas las cadenas de que $A$ es parte), de suerte que $\varphi(A_0) \subseteq \varphi(K) \subseteq L = (\varphi(A))_0$. Por otra parte, $\varphi(A) \subseteq \varphi(A_0)$ (dado que $A \subseteq A_0$) y $\varphi(\varphi(A_0)) \subseteq \varphi(A_0)$ (dado que $\varphi(A_0) \subseteq A_0$). En otras palabras, $\varphi(A_0)$ es una $\varphi$-cadena que incluye a $\varphi(A)$. Por lo tanto, $(\varphi(A))_0 \subseteq \varphi(A_0)$ (pues $(\varphi(A))_0$ es la intersección de todas las $\varphi$-cadenas que incluyen a $\varphi(A)$). Resulta, pues, que $\varphi(A_0) = (\varphi(A))_0$. Q.E.D.

14 Hoy diríamos "de modo que $A_0 \cap \Sigma \neq \varnothing$". Como Dedekind no reconoce la existencia del sistema vacío, para él un sistema existe si y sólo si contiene elementos.

El TIC es inmediatamente aplicable a la estructura $\langle\mathcal{N},\sigma,\mathbf{1}\rangle$ descrita en las pp. 150-51, en que $\sigma$: $\mathcal{N} \to \mathcal{N}$ es inyectiva y $\mathbf{1} \in \mathcal{N}$ pero $\mathbf{1} \notin \sigma(\mathcal{N})$. Estas características implican que $\langle\mathcal{N},\sigma,\mathbf{1}\rangle$ satisface los primeros cuatro axiomas de Peano. Ahora podemos ver, a la luz del TIC, que $\langle\mathcal{N},\sigma,\mathbf{1}\rangle$ satisface además el axioma **P5** si y sólo si $\mathcal{N}$ es la $\sigma$-cadena de $\mathbf{1}$ (si $\mathcal{N} = \mathbf{1}_0$, en la notación de Dedekind). En efecto, en virtud de ese teorema, $\mathbf{1}_0$ es parte de cualquier sistema $k$ si (i) $\{\mathbf{1}\} \subseteq k$ y (ii) $x \in \mathbf{1}_0 \cap k$ implica que $\sigma(x) \in k$. Ahora bien, (i) equivale a la cláusula $\mathbf{1} \in k$ ('$1 \varepsilon k$') del axioma **P5**; por otra parte, si $\mathcal{N} = \mathbf{1}_0$, (ii) dice precisamente que $\forall x((x \in \mathcal{N} \wedge x \in k) \to \sigma(x) \in k)$ ('$x \varepsilon \mathrm{N}.\ x \varepsilon k : \supset_x .\ x + 1 \varepsilon k$'). Por lo tanto, bajo las condiciones antedichas, $\mathcal{N} \subseteq k$. Por otra parte, si $\mathcal{N} \neq \mathbf{1}_0$, tenemos que $\mathcal{N} \not\subseteq \mathbf{1}_0$ aunque $\mathbf{1} \in \mathbf{1}_0$ y si $x \in \mathcal{N}$ y $x \in \mathbf{1}_0$, $\sigma(x) \in \mathbf{1}_0$. Por lo tanto, si $\mathcal{N} \neq \mathbf{1}_0$, el propio sistema $\mathbf{1}_0$ proporciona un ejemplo de una colección de objetos que reúne las condiciones prescritas a $k$ en el axioma **P5** y sin embargo no incluye a $\mathcal{N}$.

De este resultado se deduce sin dificultad la importante conclusión siguiente: Supongamos que $\sigma$: $S \to S$ es una aplicación inyectiva cualquiera y que existe un objeto $\mathbf{1} \in S$ tal que $\mathbf{1} \notin \sigma(S)$ (el lector recordará que esto sólo es posible si $S$ es un sistema infinito en el sentido de Dedekind). Entonces, si $\mathbf{1}_0$ denota la $\sigma$-cadena de $\mathbf{1}$ y entendemos que $\mathbf{1}_0$ es N, $\mathbf{1}$ es 1 y $\sigma$ es la aplicación $x \mapsto x + 1$, *la estructura* $\langle\mathbf{1}_0,\sigma,\mathbf{1}\rangle$ *satisface los axiomas* **P1**–**P5** *y puede por lo tanto prestar todos los servicios que la matemática espera de los números naturales*. En efecto, $\langle\mathbf{1}_0,\sigma,\mathbf{1}\rangle$ satisface **P1**–**P4** por la forma cómo la hemos definido y satisface **P5** en virtud del TIC, puesto que $\mathbf{1}_0$ es la $\sigma$-cadena de $\mathbf{1}_0$. Este argumento demuestra que todo sistema infinito en el sentido de Dedekind incluye un subsistema capaz de brindar un modelo de los axiomas de Peano.

En 1888 Dedekind no menciona, por cierto, los axiomas publicados por Peano en 1889, pero caracteriza el tema de la aritmética elemental como acabo de hacerlo. Dice que un sistema $N$ se llama *simplemente infinito* si hay una aplicación similar (esto es, inyectiva) $\varphi$ de $N$ en $N$ tal $N$ es la $\varphi$-cadena de un elemento suyo que no pertenece a $\varphi(N)$. Dedekind denota este elemento con el signo '1' y lo llama el *elemento básico* (*Grundelement*) de $N$. Dice también que el sistema simplemente infinito $N$ está *ordenado* por la aplicación $\varphi$. "La esencia de un sistema simplemente infinito $N$ consiste pues en la existencia de una aplicación $\varphi$ de $N$ y un elemento 1 que satisfagan las condiciones siguientes:"

α. $\varphi(N) \subseteq N$.
β. $N = 1_0$ (la φ-cadena de 1).
γ. $1 \notin \varphi(N)$.
δ. φ es similar (inyectiva).[15]

El estudio de esta esencia es la tarea propia de la aritmética, como indica la siguiente:

> DEFINICIÓN. Si en la consideración de un sistema simplemente infinito *N* ordenado por una aplicación φ, uno prescinde completamente de la particular índole de sus elementos, se fija meramente en que es posible distinguirlos y atiende sólo a las relaciones mutuas en que los pone la aplicación ordenadora φ, dichos elementos se llaman *números naturales* o *números ordinales* o simplemente *números* y el elemento básico 1 se llama el *número básico* (*Grundzahl*) de la *serie numérica N*. [ . . . ] Las relaciones o leyes que se derivan pura y exclusivamente de las condiciones α, β, γ, δ en el #71 y por ende son siempre las mismas en todos los sistemas ordenados simplemente infinitos, —cualesquiera que sean los nombres que casualmente se asignen a los elementos particulares— constituyen el objeto inmediato de la *ciencia de los números* o *aritmética*.
> (Dedekind 1888, #73)

A mi modo de ver, esta visión de la aritmética elemental como la teoría de los sistemas simplemente infinitos queda plenamente vindicada cuando Dedekind define en sus términos la relación ‘mayor que’ y las funciones aritméticas básicas: suma, multiplicación y exponenciación (de enteros positivos). Digamos con él que *u* es el número *siguiente* a *v* si *u* y *v* pertenecen al sistema simplemente infinito *N* ordenado por la aplicación φ y $u = \varphi(v)$. Dedekind demuestra que, aparte del elemento básico 1, todo elemento de *N* es el número siguiente a otro elemento de *N*.[16] Por otra parte, es fácil ver

15 Dedekind 1888, #71. He colocado el texto de los postulados α–δ fuera de la cita entre comillas porque no uso la misma notación que Dedekind.

16 En efecto, si $f\colon S \to S$ es cualquier aplicación y $A \subseteq S$, la *f*-cadena de *A*, $A_0 = A \cup f(A_0)$ (Dedekind 1888, #58). Por lo tanto, $N = 1_0 = \{1\} \cup \varphi(1_0) = 1 \cup \varphi(N)$. La citada proposición #58 se prueba así: Por la proposición #57 (demostrada en la nota 13), $f(A_0) = (f(A))_0$. Para abreviar, estipulamos que $L = f(A_0)$ y $K = A \cup L$. Entonces $f(A) \subseteq L$ (ya que $A \subseteq A_0$) y, como *L* es una *f*-cadena, la demostración del lema utilizado en la nota 13 para probar la proposición #57 indica que también *K* es una *f*-cadena tal que

que ningún $x \in N$ está contenido en la φ-cadena $(\varphi(x))_0$ del número siguiente a $x$.[17] Sea $K$ una φ-cadena cualquiera y $n \in N$. Basándose en las dos proposiciones recién citadas, Dedekind establece que (i) si $n \notin K$, $K \subseteq \varphi(n_0)$; (ii) si $n \notin K$ pero $\varphi(n) \in K$, $K = \varphi(n_0)$; (iii) hay un y sólo un $k \in K$ tal que $K = k_0$ y (iv) si $n \neq m \in N$, $n_0 \subseteq \varphi(m_0)$ si y sólo si $m_0 \not\subseteq \varphi(n_0)$.[18] Dedekind ofrece la siguiente definición: Si $n, m \in N$, $n$ es *mayor* que $m$ (simbólicamente: $n > m$), si y sólo si $n_0 \subseteq \varphi(m_0)$. En virtud de los resultados antedichos, cada par de números $n, m \in N$ satisface una y sólo una de las tres relaciones siguientes: $n > m$, o $n = m$, o $m > n$.[19] Dedekind define luego recursivamente la suma, la multiplicación y la exponenciación. Pero antes demuestra el importante Teorema de la Definición por Inducción, que asegura la univocidad de las definiciones recursivas (Dedekind 1888, #126). Mediante este teorema Dedekind demuestra asimismo que entre dos sistemas simplemente infinitos cualesquiera hay una única aplicación biyectiva que asigna al elemento básico del uno el elemento básico del otro y al siguiente de cada elemento del primer sistema el siguiente del elemento correspondiente del segundo. En virtud de ello, obviamente, todos los sistemas simplemente infinitos son intercambiables en su papel de "serie numérica" y el tema de la aritmética resulta ser precisamente la *estructura* realizada en cualquiera de

---

$A \subseteq K$ y $f(K) \subseteq L$. Como $A_0$ es la intersección de todas las $f$-cadenas que incluyen a $A$, $A_0 \subseteq K$. Por otro lado, $A \subseteq A_0$ y $L = f(A_0) \subseteq A_0$. Por lo tanto, $K = A \cup L \subseteq A_0$, de suerte que $K = A_0$.

17 En efecto, el elemento básico $1 \notin (\varphi(1))_0$, puesto que $1 \notin \varphi(N)$ y —en virtud de Dedekind 1888, #57 (demostrado en la nota 13)— $\varphi(N) = \varphi(1_0) = (\varphi(1))_0$. Supongamos ahora que la tesis vale para un cierto $n \in N$. Sea $p$ el número siguiente a $n$. Tenemos, pues, que $n \notin p_0$ y por ende difiere de cada $q \in p_0$. En consecuencia, para cada $q \in p_0$, $p = \varphi(n) \neq \varphi(q) \in \varphi(p_0) = (\varphi(p))_0$ (la desigualdad se debe a que φ es inyectiva; la última igualdad se basa en el citado #57). Por lo tanto, si la tesis vale para $n$ vale también para el número siguiente a $n$. En consecuencia, por el TIC, vale para todos los elementos de $N$. Q.E.D.

18 Las proposiciones (i), (ii), (iii) y iv) se enuncian y demuestran en Dedekind 1888, ##85, 86, 87 y 88, respectivamente.

19 Si $n = m$, $n_0 = m_0$. Como $m \in m_0$ pero $m \notin \varphi(m_0)$ —por la proposición demostrada en el nota 17—, tenemos que $n_0 \not\subseteq \varphi(m_0)$, es decir, que $n \ngtr m$. Intercambiando $n$ y $m$ en este argumento concluimos que si $m = n$, $m \ngtr n$. Por otra parte, si $n \neq m$, la proposición (iv) implica que se cumple una y sólo una de las alternativas: $n > m$, $m > n$.

ellos. Me refiero a estos asuntos en el Apéndice XI (véase también el Apéndice VIII).

### *2.3.3 Frege (1884)*

En su definición de la aritmética Dedekind adopta resueltamente lo que hoy llamamos una concepción estructuralista de las matemáticas. Desde este punto de vista cualquier sistema de cosas —no importa su origen y condición— constituye el objeto de una determinada disciplina matemática en cuanto se logra concebirlo como una realización de la especie de estructura que es el tema propio de esa disciplina. Ello permite entender que las matemáticas sean no —como se ha dicho a veces irreflexivamente— el *lenguaje* de la ciencia,[20] pero sí un componente principal del *pensamiento* científico. El estructuralismo matemático asoma ya en los escritos de Leibniz, y Newton describió en una oportunidad el espacio y el tiempo absolutos como sistemas de objetos *individualizados exclusivamente por sus relaciones mutuas*.[21] Pero fue durante el gran florecimiento de las matemáticas en el siglo XIX que el punto de vista estructuralista se fue imponiendo, sin hacer ruido, a sus practicantes. Se ha destacado con justicia la contribución de la geometría proyectiva a este proceso (Nagel 1939, Freudenthal 1974). Permítaseme mencionar sin comentarios el hecho siguiente: si $\mathcal{H}$ es un haz de rectas en el espacio ordinario —es decir, si $\mathcal{H}$ es el sistema de todas las rectas que concurren en un punto dado— y llamamos 'punto' a cada recta de $\mathcal{H}$ y 'recta' a cada plano

[20] El discurso científico se conduce en castellano, inglés, chino, etc. Algunos fragmentos sencillos se han traducido a una que otra escritura conceptual: BS, **CP1=**, etc. Pero sería grotesco pretender que las matemáticas *son* uno de estos "lenguajes".

[21] "Del mismo modo que las partes del tiempo se individualizan por el orden, de suerte que, por ejemplo, si el día de ayer pudiera permutar su orden con el día de hoy y tornarse posterior perdería su individualidad y ya no sería más el día de ayer sino el de hoy; así también las partes del espacio se individualizan por sus posiciones, de modo que si dos cualesquiera de ellas pudiesen permutar sus posiciones, permutarían a la vez su identidad y cada una se convertiría como individuo en la otra. Sólo en virtud del orden y las posiciones relativas se conciben las partes del tiempo y del espacio como siendo esas mismas que de veras son; y no tienen otro principio de individuación que ese orden y esas posiciones, las cuales, por lo tanto, no pueden cambiar" (Newton, "De Gravitatione et equipondio fluidorum", en Hall y Hall 1978, p. 103).

determinado por dos de ellas, entonces $\mathcal{H}$, así descrito, realiza todos los teoremas de la geometría proyectiva plana. Sorprende por eso que Frege, cuya tesis doctoral (1873) versó sobre un tema de geometría proyectiva, haya resistido el enfoque estructuralista aun allí donde, como en la teoría axiomática de la geometría de Hilbert, su necesidad saltaba a la vista.[22] La diferencia esencial entre su fundamentación de la aritmética y la propuesta por Dedekind se debe justamente a esta extraña resistencia —o, mejor dicho, ceguera— de Frege. Así como los teólogos demandan un concepto de Dios que no tenga más que un solo ejemplar, Frege exige una definición de número que certifique la unicidad del uno, el dos, el tres, etc. Como veremos, esta exigencia lo indujo a la contradicción que arruina su teoría.[23]

La primera contribución importante de Frege a la fundamentación de la aritmética aparece en la tercera parte de *Escritura conceptual* (1879, pp. 55–87). Consiste en la demostración de varios teoremas de lo que Frege llama "Teoría general de las series (*allgemeine Reihenlehre*)". Uno de ellos provee, como el TIC de Dedekind, una justificación directa para la inducción matemática finita. Camuflados en la escritura BS, estos resultados de Frege no llamaron la atención de Dedekind y Peano cuando elaboraban sus teorías de la aritmética. Como suele ocurrir en matemáticas, el genio del autor se

[22] Sobre los *Fundamentos de la geometría* (Hilbert 1899) Frege y Hilbert intercambian cartas —reproducidas en Frege WB, 60–76— en diciembre de 1899 y enero de 1900. La obra es objeto también de dos artículos muy críticos de Frege (1903a, b). Es pertinente asímismo la correspondencia de Frege con el joven matemático Heinrich Liebmann, hijo de su colega Otto Liebmann, y ayudante a la sazón en Göttingen, quien le había procurado un ejemplar de los apuntes —autorizados por el profesor— del curso "Elementos de geometría euclidiana" (invierno de 1898/99) en que Hilbert presentó su axiomatización. El 29 de julio de 1900 Frege devolvió este trabajo a Liebmann con una carta en la que le dice que, "aunque contiene varias invenciones ingeniosas, en conjunto lo estimo malogrado (*verfehlt*) y utilizable sólo tras mucha crítica" (WB, pp. 147s.). Resnik 1980, cap. 3, contiene una buena exposición de "La controversia entre Frege y Hilbert". Aunque, a mi modo de ver, Resnik exagera los aciertos del primero, admite que "Frege no apreció el potencial de la obra de Hilbert para la matemática en general", mostrándose "ciego" en un grado que nos resulta "chocante (*shocking*)" para la evidente analogía entre la axiomática hilbertiana y lo que él mismo había hecho con la aritmética (p. 114).

[23] En definitiva, no hay *multiples* sistemas numéricos que realicen la teoría de la aritmética de Frege porque, como ella es contradictoria, simplemente no hay *ninguno*. No faltará quien vea aquí otro paralelo con la teología.

manifiesta principalmente, no en la derivación misma de los teoremas —tarea relativamente fácil una vez bien planteada— sino en la selección de los conceptos apropiados para formularlos. Así como Dedekind da con el concepto muy simple y general de *cadena* —del que la serie numérica resulta ser un ejemplo— y demuestra el TIC para toda cadena, así Frege introduce los conceptos no específicamente aritméticos de *serie determinada por un procedimiento* y *propiedad hereditaria en una tal serie* y prueba teoremas sobre tales propiedades. Según Frege, "la inferencia de $n$ a $n + 1$" —esto es, la inducción matemática finita— se puede reducir a las leyes lógicas *únicamente* a través de estos conceptos ideados por él (Frege 1884, p. 93).[24]

Para no cansar al lector, explicaré los conceptos de Frege en castellano salpicado con fórmulas de la lógica actual.[25] Me tomo además la libertad de reemplazar los ideogramas churriguerescos que Frege usa para representarlos por otros mucho más sencillos. Numero las explicaciones para facilitar su consulta.

24 Las palabras 'se puede ... únicamente' corresponden a la frase alemana 'wird es allein möglich' empleada por Frege. Con ella expresa una apreciación singularmente exagerada de su propia obra. Como vimos, Dedekind no necesita los conceptos fregeanos para llegar al TIC con una soltura y elegancia que una persona condenada a leer únicamente a Frege no se soñaría siquiera. Claro que el TIC es un teorema de una teoría general de conjuntos, y como tal —diríamos hoy— no logra reducir la inducción matemática a "leyes lógicas universales". Pero inmediatamente después de probar el TIC en la versión conjuntista que hemos visto (Sección 2.3.2), Dedekind da un paso que —tomado al pie de la letra— reduciría en efecto la teoría de conjuntos a la lógica, *si no la hiciera inconsistente*. Me explico: El TIC enuncia ciertas condiciones que aseguran que una cadena $A_0$ está incluida en un dado sistema $\Sigma$. Por lo tanto, según Dedekind, el TIC puede invocarse para establecer que todos los elementos de $A_0$ tienen cierta propiedad $\mathfrak{E}$ (o cumplen con la condición especificada en cierta oración $\mathfrak{S}$): basta suponer que $\Sigma$ designa "el sistema de todas las cosas que poseen la propiedad $\mathfrak{E}$ (o para las cuales vale la oración $\mathfrak{S}$)" (1888, #60). Pero el supuesto de que existe, para cada propiedad $\mathfrak{E}$ (o condición $\mathfrak{S}$), el sistema de las cosas que poseen $\mathfrak{E}$ (o satisfacen $\mathfrak{S}$) implica, como sabemos, la paradoja de Russell. Gracias a Dios, para edificar la aritmética *à la* Dedekind no hace falta entender literalmente el pasaje citado: basta suponer que $\Sigma$ designa el sistema de *todos los elementos de* $A_0$ que tienen la propiedad $\mathfrak{E}$ (o cumplen con la condición $\mathfrak{S}$). Así restringido, el supuesto no implica, que yo sepa, contradicción alguna.

25 Véase el Apéndice IX. Adviértase que para explicar a Frege no puedo limitarme a utilizar el conocido cálculo de primer orden que allí se explica en detalle, sino que debo

[S0] Un ‘procedimiento’ (*Verfahren*), en el sentido de Frege, es simplemente una relación binaria. Si $R$ es un procedimiento y $Ruv$, diremos con Frege que $v$ resulta de aplicar $R$ a $u$.

En la explicaciones siguientes me refiero a un procedimiento arbitrario pero fijo que —imitando a Frege— designo con la letra minúscula $f$.

[S1] Una propiedad $F$ *es hereditaria en la serie determinada por* $f$ (simbólicamente: $\mathfrak{H}_f F$) si y sólo si $\forall u(Fu \rightarrow \forall v(fuv \rightarrow Fv))$, esto es, si y sólo si, cuando un objeto $v$ resulta de aplicar $f$ a un objeto $u$ que tiene la propiedad $F$, $v$ también tiene la propiedad $F$. En vez de ‘la serie determinada por $f$’ diré ‘la $f$-serie’ y en vez de ‘hereditaria en la $f$-serie’ diré ‘$f$-hereditaria’.

[S2] *$u$ $f$-precede a $v$* y *$v$ $f$-sucede a $u$* (simbólicamente: $u <_f v$) si y sólo si $\forall F(\mathfrak{H}_f F \rightarrow (\forall w(fuw \rightarrow Fw) \rightarrow Fv))$, esto es, si y sólo si, cualquiera que sea la propiedad $F$, $v$ tiene $F$ si $F$ es $f$-hereditaria y si el resultado de aplicar $f$ a $u$ siempre tiene $F$.

[S3] *$v$ pertenece a la $f$-serie iniciada con $u$* y *$u$ pertenece a la $f$-serie que llega a $v$* (simbólicamente: $u \leq_f v$ ) si y sólo si $u = v$ o $u <_f v$.

[S4] Supongamos ahora que $f$ es un procedimiento tal que $fuv$ y $fuw$ sólo si $v = w$. En tal caso, decimos con Frege que $f$ es un procedimiento unívoco (*eindeutig*).[26] Como puede verse, un procedimiento unívoco es simplemente una aplicación de todo el universo de objetos en sí mismo: $fxy$ si $y$ es el valor de tal aplicación en el argumento $x$.

---

ocasionalmente recurrir al cálculo de segundo orden esbozado en la sección G de dicho apéndice. Éste envuelve cuantificación sobre variables predicativas. Por ejemplo, la fórmula $\exists R \forall x \forall y(Rxy \rightarrow \neg(x = y))$ dice que hay una relación binaria $R$ que subsiste sólo entre objetos no-idénticos.

[26] Frege escribe $\underset{\delta}{\overset{\varepsilon}{\mathrm{I}}} f(\delta,\varepsilon)$ para decir que el procedimiento $f$ es unívoco. Este es el más sencillo de los ideogramas que emplea para representar los cuatro conceptos definidos arriba.

Los teoremas de la teoría de las series presentada en Frege 1879 se deducen de estas las definiciones [S1]–[S4] y los axiomas lógicos enunciados en la Capítulo 2.2 (y en el Apéndice X), mediante las reglas de inferencia mencionadas en dicho capítulo (esto es, *modus ponens* y una regla tácita de sustitución de variables). La proposición siguiente constituye el fundamento de la inducción matemática finita:

81. $(Fx \rightarrow (\mathfrak{H}_f F \rightarrow (x <_f z \rightarrow Fz)))$

(Si $F$ es una propiedad $f$-hereditaria y $x$ tiene $F$ y $z$ $f$-sucede a $x$, $z$ también tiene $F$).

Obsérvese cuánto más general que el TIC de Dedekind es este "teorema de la inducción" fregeano. Aquél es relativo a una *aplicación* arbitraria φ, pero la operación $f$ a que éste se refiere no tiene que ser una aplicación: es lisa y llanamente una relación binaria cualquiera. La demostración es facilísima, si tenemos presentes las definiciones de 'propiedad $f$-hereditaria' y '$f$-sucede'. La proposición 81 queda establecida si probamos que $Fy$ está implicada por la conjunción de las tres premisas siguientes: (i) $Fx$, (ii) $\mathfrak{H}_f F$, o sea, $\forall u(Fu \rightarrow \forall v(fuv \rightarrow Fv))$ y (iii) $x <_f y$, esto es, $\forall G(\mathfrak{H}_f G \rightarrow (\forall v(fxv \rightarrow Gv) \rightarrow Gy))$. La conjunción de (i) y (ii), combinada con el Axioma 58, implica que (iv) $\forall v(fxv \rightarrow Fv))$. La conjunción de (ii) y (iii), combinada con el Axioma 58, implica que (v) $(\forall v(fxv \rightarrow Fv) \rightarrow Fy)$.[27] La conjunción de (iv) y (v) implica $Fy$. Evidentemente, si entendemos que $x$ denota el primer número natural y $fuv$ significa '$v$ es el siguiente del número $u$', la proposición 81 combinada con la regla *modus ponens* autoriza la inducción matemática finita.

Convencido de que la aritmética sólo puede ser una ciencia si sus asertos se refieren a objetos bien determinados, Frege dedica la parte positiva de *Fundamentos de la aritmética* (1884) a establecer qué objetos son los números. No puedo examinar aquí las otras soluciones de esta cuestión que Frege

27 En esta segunda aplicación del Axioma 58, $(\forall x\Phi(x) \rightarrow \Phi(a))$, hay que entender como Frege que la variable ligada $x$ puede ser predicativa y por ende puede ser sustituida con la variable Γ que aparece ligada en la premisa (iii). Frege nos pediría sustituir en $(\forall x\Phi(x) \rightarrow \Phi(a))$, $x$ por Γ, $a$ por Φ, y Φ(Γ) por $(\mathfrak{H}_f\Gamma \rightarrow (\forall v(fxv \rightarrow \Gamma v) \rightarrow \Gamma y))$; cf. p. 131, nota 6.

fustiga en la parte inicial —negativa— de su libro, aunque la crítica devastadora a que las somete contribuye decididamente a motivar la suya. Los objetos que propone —si existiesen de veras— constituirían, como veremos, la serie determinada por cierto procedimiento, de modo que la teoría de 1879 les sería aplicable.

Frege procede a averiguar qué son los números examinando el modo como corrientemente se habla de ellos. Para eso disponemos de palabras (y signos) —*uno*, *dos*, *tres*, *cuatro* . . . (1, 2, 3, 4, . . . )— que llamaré *numerales*. También tenemos reglas precisas para construir nuevos numerales cada vez que sea menester. Ahora bien, en la conversación cotidiana los numerales funcionan como atributos o predicados —'estas *tres* gallinas pusieron ayer *cinco* huevos', 'la semana tiene *siete* días', 'hay *ochenta mil* soldados paflagonios en Moronesia'. Por esto, los manuales de gramática clasifican a los numerales como adjetivos. Sin embargo, en la fraseología matemática se comportan como sustantivos, según puede verse en los siguientes ejemplos: '*cuatro* es el cuadrado de *dos*', '*once* es el único número primo mayor que *siete* y menor que *trece*', 'el *dos*, el *tres* y el *siete* son los únicos divisores primos del *42*'. Frege dice que le interesa concebir la noción de número en la forma en que es utilizable por la ciencia, así que no se dejará turbar por el hecho de que en la vida diaria los numerales también se usan atributivamente (1884, p. 69). Además, cualquier expresión en que un numeral se usa de este modo es fácilmente sustituible por otra equivalente en que se lo emplea como sustantivo; así, en vez de 'la semana tiene siete días' siempre puedo decir —si no me importa parecer afectado— 'el número de los días de la semana es igual a siete'. No obstante esta advertencia, Frege aborda su tema con una pregunta que concierne directamente al uso atributivo (y predicativo) de los numerales, como quiera que se lo disimule con transformaciones sintácticas. Cuando empleamos un numeral como adjetivo ¿a qué —o a quién— le atribuimos una propiedad? La respuesta parece obvia: la propiedad significada mediante un adjetivo se atribuye (en castellano) al objeto nombrado por el sustantivo que dicho adjetivo califica. Así, el aserto sobre las fuerzas armadas de Paflagonia atribuye a *los soldados de esa nación estacionados en Moronesia* la propiedad de ser *ochenta mil*. Frege descarta esta respuesta debido a que la presencia militar paflagonia en Moronesia también se puede describir correctamente como *un* ejército, *cinco* divisiones, *cuarenta* regimientos, *ciento sesenta* batallones. La propiedad expresada

por el numeral ‘ochenta mil’ no se atribuye, pues, a la *realidad concreta* nombrada en nuestro aserto, sino al *concepto* de soldado paflagonio en Moronesia mediante el cual se refiere a ella. Este argumento fregeano no tiene en cuenta, al parecer, que aunque una misma realidad exhiba distintas propiedades según el concepto adoptado para objetivarla, ellas son propiedades del objeto captado, no del concepto bajo el cual se lo capta. Cuando leo *Las Bacantes* leo a la vez una tragedia *antigua* y una edición *moderna*; un escrito *extraordinario* y un impreso *muy ordinario*; una historia *simple*, pero un texto *difícil*. Sería ridículo decir que esta última aseveración atribuye dificultad no a la obra misma que estoy leyendo ahora —*Las Bacantes* de Eurípides, en la edición de E. R. Dodds— sino a un concepto general bajo el cual se la puede clasificar. Pero Frege tiene otra razón para sostener que los adjetivos numerales expresan propiedades de conceptos. Imaginemos un cuadro estadístico, por países, de los enfermos de cólera diagnosticados en 1991. Frente al nombre ‘Perú’ el cuadro registra un número de cinco cifras, pero frente a ‘Finlandia’ hay un cero. Evidentemente, el cuadro no atribuye la propiedad de ser cero al objeto descrito por la frase ‘enfermo de cólera diagnosticado en Finlandia en 1991’. De acuerdo con lo que el propio cuadro nos informa, tal objeto *no existe* y no puede, por lo tanto, tener propiedad alguna. Pero el cuadro sí podría atribuir dicha propiedad al concepto de *enfermo de cólera diagnosticado en Finlandia en 1991* y —según la tesis de Frege— de hecho se la atribuye.[28]

[28] He aquí una traducción del pasaje de Frege que he estado parafraseando:

> Si con respecto a un mismo fenómeno externo puedo decir con la misma verdad: “esto es un grupo de árboles’ y “esto son cinco árboles”, o “aquí hay cinco compañías” y “aquí hay 500 hombres”, no varía en tales casos ni lo individual, ni el todo, el agregado, sino la denominación que le doy. Pero esto indica sólo la sustitución de un concepto por otro. Esto sugiere como respuesta a la primera pregunta del párrafo anterior [“¿de quién se asevera algo mediante una indicación numérica (*Zahlangabe*)?”] que la indicación numérica contiene un aseveración sobre un concepto. Esto se percibe tal vez con máxima claridad en el caso del número 0. Si digo “Venus tiene 0 lunas”, no hay ninguna luna o colección de lunas de la que pudiera estar diciendo algo; pero le atribuyo una propiedad al *concepto* “luna de Venus”, a saber, la de no comprender nada (*nichts unter sich zu befassen*). Si digo “el coche del emperador va tirado por cuatro caballos” le atribuyo el número cuatro al concepto “caballo que tira el coche del emperador”.
>
> (Frege 1884, p. 58)

Supondré pues, con Frege, que atribuyo cierta propiedad al concepto *C* cuando asevero que los *C* son un cierto número *n*. ¿Qué propiedad es ésa? Tentativamente, Frege propone la siguiente definición recursiva (1884, p. 67):

(Δα) Al concepto *C* le corresponde (*zukommt*) el número 0 si y sólo si, cualquiera que sea el objeto *x*, *x* no es *C*.

(Δβ) Al concepto *C* le corresponde el número 1 si y sólo si hay un objeto *a* tal que *a* es *C* y cualquier objeto *b* que sea *C* es idéntico a *a*.

(Δγ) Al concepto *C* le corresponde el número *n* + 1 si y sólo si hay un objeto *a* tal que *a* es *C* y al concepto *C** expresado por la expresión (insaturada) '*x* es *C* y $x \neq a$' le corresponde el número *n*.

Frege rechaza esta propuesta en el acto, aunque admite que es tan natural que el rechazo demanda una explicación. Su principal defecto consiste, según él, en que se ha definido la expresión 'a *C* le corresponde el número *n* + 1' mediante la expresión 'a *C** le corresponde el número *n*', cuyo significado se desconoce. Utilizando las definiciones ofrecidas,

> podemos decir lo que significa (*bedeutet*) "al concepto *F* le corresponde el número 1 + 1" y luego, mediante este resultado, indicar el sentido (*Sinn*) de la expresión "al concepto *F* le corresponde el número 1 + 1 + 1", etc. Mas —para dar un ejemplo craso— con nuestras definiciones no podremos decidir nunca si a un concepto le corresponde el número *Julio César*—si el conocido conquistador de Galia es o no es un número. Además, no es posible demostrar, con ayuda de las definiciones ensayadas, que *a* tiene que ser igual a *b* si al concepto *F* le corresponde el número *a* y a ese mismo concepto le corresponde el número *b*. Por lo tanto, la expresión '*el* número que corresponde al concepto *F*' no podría justificarse y sería, por ende, generalmente imposible demostrar una identidad numérica, ya que no podríamos captar un determinado número.
>
> (Frege 1884, p. 68)[29]

Como sabemos, el Teorema de la Definición por Inducción de Dedekind (Sección 2.3.2 y Apéndice XI) asegura la univocidad de las definiciones

29 La primera oración de este pasaje muestra que en 1884 Frege todavía utilizaba —como en alemán corriente— *Bedeutung* ('significado') y *Sinn* ('sentido') como sinónimos.

recursivas y por lo tanto desarma la objeción general contra tales definiciones implícita en este texto de Frege. Pero dicho teorema presupone la existencia de una serie numérica y no se lo puede invocar para caracterizarla. Por otra parte, el texto citado nos sugiere por qué Frege no supo hallar, como Dedekind y Peano, una caracterización flexible y viable de los números, y dio en cambio con un dogma rígido y contradictorio. Frege no entendió nunca que una teoría matemática se interesa por sus objetos sólo en cuanto mantienen ciertas relaciones mutuas, y no en cuanto a lo que cada uno pudiera ser de suyo. Julio César será ciertamente un número, si integra con otros objetos lo que Dedekind llama un sistema simplemente infinito (para lo cual el difunto dictador romano es perfectamente idóneo). Pero a la aritmética no va a interesarle su programa político, ni su estilo literario, ni sus hazañas militares, sino sólo qué objeto le sigue y qué objetos le preceden en ese sistema (si él mismo no es el primero). Por eso, en su papel de número nadie lo nombraría con su nombre de familia, sino con el numeral correspondiente (como hacemos con los dedos, cuando los usamos para sacar cuentas). Incapaz de ver estas cosas que a muchos nos resultan tan obvias, Frege censura a Hilbert por modelar el plano euclidiano en un sistema de pares de números reales.[30] La misma ceguera motiva, me parece, su afán de encontrar, para cada número natural, un objeto único que se identifique irremplazablemente con él.

30 Hilbert llama Ω al sistema de los números algebraicos, esto es, los números reales que son soluciones de ecuaciones polinomiales con coeficientes enteros. Frege, 1903b, p. 374, escribe:

> En la p. 20 [de Hilbert 1899] se dice: "Tomamos como punto un par de números $\langle x,y \rangle$ del dominio Ω", etc. Si mediante la definición y los axiomas pertinentes se le hubiese ya conferido una denotación (*Bedeutung*) a la palabra 'punto', no se podría hacerlo aquí otra vez. La cosa hay que pensarla así quizás: el concepto *es un par de números del dominio* Ω, que es de primer orden, igual que el concepto euclidiano de punto, debe caer bajo el concepto hilbertiano de segundo orden (si es que hay tal concepto). Molesta, claro está, el uso de la palabra 'punto' en ambos casos, pues evidentemente tiene en cada uno una denotación diferente.

El 27 de diciembre de 1899 Frege ya le había objetado por carta a Hilbert que sus conceptos de 'punto' y 'entre' no son unívocos (WB, p. 63). Hilbert le envió a vuelta de correo la declaración sobre "el sistema *amor*, *ley*, *deshollinador*, . . . " que cité al comienzo de la Sección 1.8.1.

Inmediatamente después de criticar como hemos visto la definición recursiva propuesta, Frege le objeta que ya la base de la recursión ($\Delta\alpha$) nos dice solamente lo que significa 'corresponder (a un concepto) el número 0', mas no lo que el número 0 *es*. Esto quiere decir que, en el mejor de los casos, la definición determinaría a lo sumo el uso correcto de los *adjetivos* numerales, mas no el significado de los *sustantivos* numerales que nombran a los objetos de la aritmética. Para obviar esta dificultad Frege intenta una maniobra familiar en matemáticas. En el lenguaje de hoy, la describiríamos sumariamente así: Frege define cierta relación de equivalencia entre pares de conceptos e identífica cada número con una de las clases de equivalencia determinadas por dicha relación. La equivalencia definida por Frege es la relación que hay entre dos conceptos a los que corresponde el mismo número. Espontáneamente, diríamos que un mismo número corresponde a dos conceptos $F$ y $G$ si el número $n$ correspondiente a $F$ es idéntico al número $m$ correspondiente a $G$. Pero este camino le estaba vedado a Frege mientras no supiera lo que significa la expresión 'el número $n$ corresponde al concepto $F$'. Su magistral ocurrencia consistió en definir la relación diádica '$X$ tiene el mismo número que $Y$' *sin haber definido aún* la propiedad '$X$ tiene un cierto número', y usar la definición de *aquella* para fijar la de *ésta*. He aquí su definición (Frege 1884, p. 85):

[$N_{\approx}$] La expresión 'el concepto $F$ es equinumeroso [*gleichzahlig*] con el concepto $G$' significa lo mismo que la expresión 'hay una relación $\phi$ que coordina biunívocamente [*beiderseits eindeutig zuordnet*] los objetos que caen bajo el concepto $F$ con los objetos que caen bajo el concepto $G$'.

Para entender esta definición, presuponemos con Frege que, si $F$ es cualquier concepto, los objetos que caen bajo $F$ forman un colectivo. Siguiendo la tradición, Frege lo llama 'la extensión (*Umfang*) de $F$'. Frege explica que una relación (diádica) $\phi$ *coordina* un objeto $a$ con otro objeto $b$ si el par $\langle a,b\rangle$ cae bajo el concepto $\phi$.[31] $\phi$ determina una coordinación biunívoca si

[31] Usando las ideas de Frege 1891, diríamos que $\phi$ coordina $a$ con $b$ si y sólo si la función $\phi$ toma el valor **v** en el argumento $\langle a,b\rangle$. Como Frege (1884, p. 84) explica prolijamente, es enteramente legítimo decir que $\phi$ coordina los objetos que caen bajo $F$

cumple los dos requisitos siguientes: (i) si el par $\langle a,b_1 \rangle$ cae bajo el concepto $\phi$, entonces $\langle a,b_2 \rangle$ cae bajo $\phi$ sólo si $b_1 = b_2$; (ii) si el par $\langle a_1,b \rangle$ cae bajo el concepto $\phi$, entonces $\langle a_2,b \rangle$ cae bajo $\phi$ sólo si $a_1 = a_2$. Utilizaré la abreviatura '$F \approx G$' para significar que el concepto $F$ es equinumeroso con el concepto $G$. En nuestra jerga actual diríamos pues que, con arreglo a [$N_\approx$], $F \approx G$ si y sólo si existe una aplicación biyectiva de la extensión de $F$ en la extensión de $G$. La relación de equinumerosidad es evidentemente una equivalencia puesto que (i) cada concepto es equinumeroso consigo mismo, (ii) si $F \approx G$, $G \approx F$, y (iii) si $F \approx G$ y $G \approx H$, $F \approx H$. Como Frege supone tácitamente que hay una suerte de reino formado por la totalidad de los conceptos, para él es claro que la relación de equinumerosidad determina una partición† de ese reino en colecciones de conceptos equinumerosos. Ellas son las clases de equivalencia determinadas por la relación. Bajo este supuesto tenemos, pues, que en virtud de [$N_\approx$] cada concepto $F$ pertenece a una y sólo una de esas clases. Ella es la extensión del concepto que expresamos con el predicado 'es equinumeroso con $F$'. Definimos con Frege (1884, p. 85):

[$N_F$] El número correspondiente al concepto $F$ es la extensión del concepto expresado mediante la fórmula '$X \approx F$'.

[$N_n$] La expresión '$n$ es un número' significa lo mismo que la expresión 'existe un concepto $F$ tal que $n$ es el número correspondiente a $F$'.

Con estas definiciones se habría vindicado finalmente —bajo los supuestos indicados— el uso de los numerales como sustantivos. Un numeral nombra un número, que según [$N_F$] es un objeto: la extensión de un concepto. Pero también se ha justificado su uso como adjetivos: el concepto cuya extensión nombra un cierto numeral se determina fijando uno de los correlatos de una relación binaria entre conceptos: la relación de equinumerosidad. Dicho numeral puede emplearse por eso sin mayor riesgo de confusión para expresar la propiedad que pertenece a un concepto cualquiera si y sólo si es equinumeroso con el correlato fijado.

---

con los objetos que caen bajo $G$ —simbólicamente: $\forall x(Fx \rightarrow \exists y(Gy \wedge \phi xy))$— aunque la extensión de $F$ esté vacía. En efecto, si $\forall x(Fx \leftrightarrow x \neq x)$, es claro que $\forall x \neg Fx$, de modo que $\forall x(\neg Fx \vee \exists y(Gy \wedge \phi xy))$. La última oración equivale a $\forall x(Fx \rightarrow \exists y(Gy \wedge \phi xy))$.

Frege asigna ahora denotaciones precisas a los numerales 0 y 1:

[$N_0$] 0 es el número correspondiente al concepto expresado mediante la fórmula '$x \neq x$'.

[$N_1$] 1 es el número correspondiente al concepto expresado mediante la fórmula '$x = 0$'.

La identificación fregeana de los objetos denotados por los demás numerales depende esencialmente de la relación binaria que Frege llama *seguir inmediatamente en la serie natural de los números*. Esta relación se define así (Frege 1884, p. 89):

[$N_\sigma$] La oración »*n* sigue inmediatamente a *m* en la serie natural de los números« (»*n folgt in der natürlichen Zahlenreihe unmittelbar auf m*«*)* significa lo mismo que la oración »hay un concepto *F* y un objeto *a* que cae bajo *F*, y *n* es el número correspondiente a *F* y *m* es el número correspondiente al concepto expresado mediante la fórmula '$Fx$ pero $x \neq a$'«.

En vez de '*n* sigue inmediatamente a *m* en la serie natural de números' escribiré '$\sigma mn$' (léase: '*n* sigue inmediatamente a *m*' o '*n* es el siguiente de *m*' o '*m* precede inmediatamente a *n*' o '*m* es el predecesor inmediato de *n*').[32] Comprobamos en el acto que, conforme a las definiciones de Frege, 1 es en efecto el siguiente de 0. Según [$N_\sigma$], $\sigma 01$ equivale a esta doble generalización existencial: hay un concepto *F* y un objeto *a* que cae bajo *F*, 1 es el número correspondiente a *F* y 0 es el número correspondiente al concepto

32 No escapará al lector que, conforme a [$N_\sigma$], un número puede ser su propio predecesor inmediato. Por ejemplo, si *n* es el número correspondiente al concepto 'punto del círculo con centro *P* y radio *r*' y *m* es el número correspondiente al concepto 'punto excéntrico del círculo con centro *P* y radio *r*', [$N_\sigma$] implica que $\sigma mn$ aunque $m = n$ (pues ambos conceptos son equinumerosos). Por lo tanto, contra lo que parecen creer algunos traductores, la serie natural de los números, en el sentido de Frege, no es lo mismo que la serie de los números naturales, en el sentido corriente. *Nuestros* números naturales son lo que Frege (1884, p. 96) llama *números finitos*, esto es, los números pertenecientes a la σ-serie iniciada con 0 (cf. [S3]). Por cierto, un número tal que $0 \leq_\sigma x$, satisface la condición $\neg\sigma xx$.

expresado mediante la fórmula ‘$Fx \wedge x \neq a$’. Dicha generalización existencial es verdadera si tiene por lo menos un ejemplo particular. Helo aquí: sea *a* el número 0 y *F* el concepto que expresamos mediante la fórmula ‘$x = 0$’. 1 es el número correspondiente a este concepto y 0 cae bajo el mismo; además 0 es precisamente el número correspondiente al concepto expresado mediante la fórmula ‘$x = 0 \wedge x \neq 0$’.

Es fácil comprobar que, si $\sigma mn$, no hay otro número $p \neq n$ tal que $\sigma mp$: si el siguiente de un número existe, es único. La relación binaria σ constituye pues lo que Frege (1879) llamaba un *procedimiento unívoco* (véase la explicación [S4] en la p. 162). También se comprueba fácilmente que, si $\sigma mn$, no hay otro número $q \neq m$ tal que $\sigma qn$: si el predecesor inmediato de un número existe, es único. Me parece, pues, razonable decir que σ es un procedimiento *biunívoco*. Resumo este resultado en una proposición que llamaré **F3** para subrayar su homología con el axioma **P3** de Peano:

[**F3**] Si $\sigma mn$ y $\sigma pq$, entonces $m = p \leftrightarrow n = q$.

Nuestras definiciones implican que 0 *no es es el siguiente de ningún número*,[33] lo cual se puede enunciar así:

[**F4**] No hay un número *m* tal que $\sigma m0$.

Un objeto *x* es un *número finito* si y sólo si *x* es un número perteneciente a la σ-serie iniciada por 0 (simbólicamente: $\forall x(Nxy \leftrightarrow 0 \leq_\sigma x)$). Obsérvese que esta propiedad es σ-hereditaria: $\forall x \forall y(Nx \wedge \sigma xy \rightarrow Ny)$.[34]

33 En efecto, según [$N_\sigma$], $\sigma m0$ implicaría que 0 es el número correspondiente a un concepto *F* bajo el cual cae un objeto *a*. Tal concepto *F* no podría ser equinumeroso con el concepto *H* definido por la condición $Hx \leftrightarrow x \neq x$. Mas 0 es, por definición, el número correspondiente a un cierto concepto *F* si y sólo si $F \approx H$.

34 Como σ es un procedimiento unívoco, esta aseveración es una consecuencia inmediata de Frege 1879, Teorema 131. Pero también podemos justificarla directamente así: Recordemos las definiciones [S2] y [S3]. Sea *m* un número finito y *n* el siguiente de *m*. Por [$N_\sigma$], *n* es un número. Si $0 = m$, $1 = n$ y es claro que $0 <_\sigma n$, es decir, que *n* posee todas las propiedades σ-hereditarias del siguiente de 0, puesto que éste es el mismo *n*. Si $0 <_\sigma m$, *m* posee todas las propiedades σ-hereditarias del siguiente de 0 y, por cierto, se las trasmite a su propio siguiente *n*, de modo que $0 <_\sigma n$. En cualquier caso, $0 \leq_\sigma n$: *n* es un número finito.

Como $0 = 0 \rightarrow N0$, es claro que

[**F1**] 0 es un número finito.

Como $0 <_\sigma x \rightarrow Nx$, tenemos además que (por [S3] y Frege 1879, Teorema 81, demostrado arriba en la p. 163):

[**F5**] *Cada* número finito posee *todas* la propiedades σ-hereditarias del 0.

Así pues, la justificación del razonamiento inductivo en aritmética fluye directamente de la definición fregeana de número finito. En cambio, la proposición siguiente es más difícil de probar:

[**F2**] Si *m* es un número finito cualquiera, existe un número finito *n* tal que $\sigma mn$.

La prueba, meramente esbozada por Frege en 1884, se basa en la simple observación siguiente: Si *m* es un número finito, hay exactamente *m* números finitos que σ-preceden a *m*. Por lo tanto, si *m* existe, también existe el siguiente de *m*: éste es el número correspondiente al concepto 'número finito perteneciente a la σ-serie que llega a *m*' (o sea, 'número *x* tal que, o bien $0 \leq_\sigma x <_\sigma m$, o bien $x = m$'). Como Frege (1884, p. 95) nos advierte, esta observación presupone que ningún número finito se σ-sucede a sí mismo, esto es, que $\forall x(Nx \rightarrow \neg(x <_\sigma x))$.[35] No puedo detenerme a demostrarlo.[36]

[35] Supongamos, por el contrario, que $m <_\sigma m$. En tal caso, el concepto 'número finito perteneciente a la σ-serie que llega al predecesor inmediato de *m*' sería equinumeroso con el concepto 'número finito perteneciente a la σ-serie que llega a *m*', y el número correspondiente a este último concepto no podría ser el siguiente de *m*..

[36] La proposición $\forall x(Nx \rightarrow \neg(x <_\sigma x))$ es el Lema 52 en Wright 1983, Sección XIX. Wright bosqueja una demostración en las pp. 165–67. Ella depende de una definición y dos sublemas. Defino con Wright: el concepto *F* es menos amplio que el concepto *G* ($F \prec G$) si y sólo si hay un concepto *H* tal que $F \approx G \wedge H$ y $\exists x(Gx \wedge \neg Hx)$. (Aquí '$G \wedge H$' designa el concepto bajo el cual caen precisamente todos los objetos que caen bajo *G* y bajo *H*). Los sublemas demostrados por Wright se pueden enunciar así: [521] Si $u <_\sigma v$ y *u* es el número correspondiente a un concepto *U* y *v* es el número correspondiente a un concepto *V*, entonces $U \prec V$. [522] Si el número correspondiente al concepto *F* es un número finito, entonces $\neg(F \prec F)$ (*F* no es menos amplio que sí

Consideremos, pues, un número finito $m$. Sea $S_m$ el concepto definido por la condición: $S_m x \leftrightarrow Nx \wedge x \leq_\sigma m$. Denotaré con $\nu_m$ el número correspondiente a $S_m$. Como $N$ es σ-hereditaria, es claro que $\sigma m \nu_m$ implica que $N\nu_m$ ($\nu_m$ es un número finito). Así, para demostrar [**F2**] bastaría establecer que $\sigma m \nu_m$. Para ello, conviene prestar atención a la propiedad $S$ definida por: $Sx$ si y sólo si $x$ es un número finito tal que $\sigma x \nu_x$. (En otras palabras: el número finito $x$ tiene la propiedad $S$ si y sólo si el siguiente de $x$ es el número correspondiente al concepto $S_x$). Si se demuestra (i) que $S0$ y (ii) que $S$ es σ-hereditaria, **F5** implica que $\forall x(Nx \rightarrow Sx)$. Es fácil probar (i). En efecto, $S0 \leftrightarrow \nu_0 = 1$, esto es, si y sólo si hay un objeto único $x$ tal que $S_0 x$. Como es imposible que exista un objeto $x$ tal que $x <_\sigma 0$,[37] tenemos que $\forall x(x \leq_\sigma 0 \leftrightarrow x = 0)$, lo cual implica que 0 es el único objeto que tiene la propiedad $S_0$, de modo que efectivamente $\nu_0 = 1$. Probar (ii) no es tan fácil. Sea $m$ un número finito cualquiera. Entonces, $Sm \leftrightarrow \sigma m \nu_m$. Hay que probar que $Sm \rightarrow S\nu_m$. Para simplificar la tipografía, pongo $\nu_m = n$. Sea $p$ el número correspondiente al concepto $T_n$, definido por la condición $T_n x \leftrightarrow S_n x \wedge x \neq n$. En virtud de [$N_\sigma$], es claro que $\sigma p \nu_n$. Como $S_n x \leftrightarrow Nx \wedge (x = n \vee x <_\sigma n)$, tenemos que $T_n x \leftrightarrow Nx \wedge x <_\sigma n$. Por hipótesis, $\sigma mn$. Es claro, entonces, que $T_n m$.[38] Habría que probar, además, que $T_n x \wedge x \neq m \rightarrow x <_\sigma m$.[39] Si lo damos por

mismo). Con los supuestos de Wright es fácil demostrar además que si $x$ es un número finito, $x$ es el número correspondiente a algún concepto. Combinados con ella, 521 y 522 implican que ningún número finito se σ-precede a sí mismo.

37 Sea $F$ la propiedad definida por: $Fx$ si y sólo si hay un concepto $G$ y un objeto $u$ tales que $Gu$ y $x$ es el número correspondiente a $G$. Si $w$ es el siguiente de cualquier número, [$N_\sigma$] implica que $Fw$. Así pues, trivialmente, $F$ es una propiedad σ-hereditaria. Por lo tanto, si $u$ es un número cualquiera, $F$ es una de las propiedades σ-hereditarias del siguiente de $u$. Como obviamente, $\neg F0$, resulta que $\neg(u <_\sigma 0)$: 0 no posee todas las propiedades σ-hereditarias del siguiente de $u$.

38 Según lo que recién hemos visto, $T_n m \leftrightarrow Nm \wedge m <_\sigma n$. $Nm$ por hipótesis y $m <_\sigma n$ si y sólo si $n$ tiene todas las propiedades σ-hereditarias que posee el siguiente de $m$. Esta es una verdad trivial si, como hemos supuesto, $n$ es justamente el siguiente de $m$.

39 $T_n x \wedge x \neq m \rightarrow x <_\sigma m$ es una consecuencia directa de la conjunción de las dos aseveraciones precedentes con el Lema 5121 de Wright 1983, Sección XIX, cuya demostración Wright esboza en las pp. 163–65. Ella depende de la tricotomía de los números finitos: $\forall x \forall y(Nx \wedge Ny \rightarrow x <_\sigma y \vee x = y \vee y <_\sigma x)$ (Lema 51212 de Wright) y de la Proposición 124 de Frege 1879: Si $f$ es un procedimiento unívoco y $fxy$, entonces $\forall z(y <_f z \rightarrow y \leq_f z)$ (Lema 51211 de Wright).

hecho, concluimos que $T_n x \leftrightarrow x \leq_\sigma m \leftrightarrow S_m x$, de suerte que $T_n \approx S_m$. Por consiguiente, $p = \nu_m = n$ y $\sigma n \nu_n$. Así queda establecido que $Sm \rightarrow Sn$: $S$ es, pues, σ-hereditaria. En virtud de (i) y (ii), cada número finito $m$ cumple la condición $Sm$ y, por ende, hay un número que le sigue inmediatamente en la serie natural de los números.

Así pues, con sólo llenar las lagunas del razonamiento anterior sería posible deducir los cinco Axiomas de Peano de leyes lógicas suplementadas con las definiciones fregeanas [$N_\approx$], [$N_F$], etc. Desgraciadamente, en el sistema de Frege, estas definiciones sólo pueden enunciarse como tales bajo un supuesto que, como sabemos, implica la paradoja de Russell (Capítulo 1.6):[40] si $F$ es un concepto bien definido, existe la extensión de $F$. Arriba vimos que este supuesto entra en la definición [$N_\approx$], y sobre todo en la definición [$N_F$], de la que penden la objetividad y unicidad de los números fregeanos. En el Apéndice XII explico mejor de qué modo ese supuesto está presente en el sistema maduro de Frege, y lo hace contradictorio. Aquí sólo me interesa destacar que si no hubiera contradicción o si un neo-fregeano pudiera eliminarla, los argumentos de Frege establecerían a lo sumo que esas extensiones de conceptos que él llama *números finitos* forman *un* modelo de la especie de estructura caracterizada por los Axiomas de Peano. Por lo tanto,

40 Wright (1983) buscó esquivar la catástrofe reemplazando la definición [$N_\approx$] por un axioma del mismo tenor. Elaborando esta idea, Boolos (1987) muestra que la aritmética puede fundarse entera sobre los principios de la lógica de segundo orden de la *Begriffschrift* (Frege 1879) y un solo axioma adicional, que puede parafrasearse así: 'Para todo concepto $F$ existe un único objeto $x$ tal que, para todo concepto $G$, $x$ es la extensión de $F$ si y sólo si hay una relación $\phi$ que coordina biunívocamente los objetos que caen bajo el concepto $F$ con los objetos que caen bajo el concepto $G$'. Boolos ofrece además una prueba de consistencia del sistema deductivo así formado, relativa a la consistencia de la teoría de conjuntos de Zermelo-Fraenkel (ZF), y también a la consistencia del sistema, estrictamente más débil que ZF, que Shoenfield (1967, § 8.5) llama "aritmética de segundo orden". Ni el axioma de Wright, ni el de Boolos pueden pasar por verdades lógicas (Boolos subraya que el suyo, combinado con los susodichos principios lógicos, implica la existencia de infinitos objetos; *vide* Demopoulos 1995, p. 231). Por este camino no se llega, pues, a vindicar la pretensión de Frege de que la matemática es parte de la lógica. Pero sí se pone de manifiesto —como destaca Boolos— el enorme logro simplificador y elucidatorio del análisis fregeano del concepto de número, "mediante el cual podemos ver cómo un vasto cuerpo de matemáticas puede deducirse de un único principio simple y obviamente consistente" (Boolos 1987, en Demopoulos 1995, p. 232).

si tales extensiones existieran, podrían legítimamente reclamar el nombre de ‘números’, mas no con mejor derecho que los elementos de cualquier otro modelo de esos axiomas.

## 2.4 LA TEORÍA DE LOS TIPOS LÓGICOS

Russell comunicó a Frege en junio de 1902 la paradoja de la clase de todas las clases que no son miembros de sí mismas.[1] Trabajaba entonces en el libro *The Principles of Mathematics,* que apareció en 1903. Aunque lo concibió —y en buena parte escribió— antes de leer a Frege, la doctrina sustentada en él concuerda en buena medida con la del pensador alemán,[2] y resulta igualmente afectada por la paradoja. Las Partes II–VII del libro se dedican a demostrar "que toda la matemática pura se ocupa exclusivamente con conceptos definibles en términos de un número muy pequeño de conceptos lógicos fundamentales, y que todas sus proposiciones se pueden deducir de un número muy pequeño de principios lógicos fundamentales", mientras que la Parte I acomete la "tarea puramente filosófica" de elucidar "los conceptos

[1] Véase la nota 22 en la p. 143. Es razonable pensar que Russell comunicó la paradoja a Frege poco después de descubrirla. Tal era, por ejemplo, la opinión de Coffa (1991, p. 114). Pero desde que se sabe que Zermelo descubrió la paradoja independientemente en 1901 (cf. la nota 3 en la p. 51), los expertos en Russell se han empeñado en adelantar la fecha en que éste dio con ella. Rodríguez Consuegra me asegura —en comunicación privada— que "fue en 1901, probablemente en mayo, aunque hay quien sostiene que en diciembre de 1900". Por otra parte, según el propio Rodríguez Consuegra (1989, p. 134), los pasajes de *The Principles* que suponen un conocimiento de la paradoja —esto es, el Capítulo 10, la versión final del Capítulo 43 y los apéndices sobre Frege y sobre la teoría de los tipos— fueron despachados a la imprenta entre junio y noviembre de 1902.

[2] En el Apéndice A del libro citado Russell expone "Las doctrinas lógicas y aritméticas de Frege", según él las veía en 1903, y señala algunas de las diferencias que lo separan de ellas. Tiles 1991, Cap. 3, muestra que, aunque coinciden en la superficie, las ideas de Frege y de Russell sobre esta materia responden a motivaciones muy diferentes. Coffa (1980, pp. 247s.) y otros han destacado que según Russell las verdades aritméticas son sintéticas, mientras que según Frege son analíticas. Pero el pasaje en que Russell introduce esta tesis (1903, § 434) remite a Russell 1900, § 11, donde se define un *juicio analítico* como "aquel cuyo predicado está contenido en el sujeto", noción ésta francamente más estrecha que la de Frege, para quien, como sabemos, una aseveración es analítica cuando se deduce de definiciones y las leyes de la lógica (cf. p. 130, nota 4).

fundamentales que la matemática acepta como indefinibles" (Russell 1903, Prefacio; 2ª ed., 1937, p. xv). La paradoja es el tema del capítulo 10, "The Contradiction". En ese mismo capítulo y en el Apéndice B, Russell esboza un método para resolverla. Se trata de la primera versión, todavía rudimentaria, de la teoría de los tipos lógicos adoptada en *Principia Mathematica* (1910–13), la obra monumental que Russell escribe en colaboración con Whitehead. En el citado Prefacio, Russell en cierto modo anuncia esta obra al declarar que lo que en 1903 ofrece al público es sólo un primer tomo, escrito informalmente en inglés y dirigido principalmente a los filósofos, y que irá seguido de un segundo tomo, formal, en escritura conceptual, para cuya redacción ha reclutado la ayuda del matemático Whitehead. Pero los *tres* tomos que finalmente escribieron juntos se alejan bastante del libro de 1903 y no se presentan como una continuación suya.

No pretendo examinar aquí la filosofía matemática de Russell.[3] Menos aún puedo evaluar en este espacio el aporte de Whitehead y Russell a un mejor entendimiento de la arquitectura de la matemática clásica y su influencia efectiva en el desarrollo ulterior de la misma.[4] Me limitaré a hablar de la teoría de los tipos, que es seguramente —después de la paradoja— la contribución más original y característica de Russell a nuestro tema. Quizás sería mejor considerarla no como una *teoría* —descripción o explicación de algún género de realidades— sino como una *preceptiva gramatical*, dirigida a impedir que se escriban expresiones capaces de generar contradicciones como la paradoja de Russell. Pero sólo mucho más tarde —respondiendo a la crítica

3 Russell mismo la explica admirablemente en su forma madura en *Introduction to Mathematical Philosophy* (1919), de la que hay varias traducciones al castellano. El lector interesado en sus comienzos, antecedentes y motivaciones puede ahora recurrir al excelente libro de Rodríguez Consuegra (1991).

4 Sospecho que, salvo por la saludable costumbre de formular las proposiciones y razonamientos matemáticos en un fragmento del lenguaje "natural" (castellano, inglés, etc.) que se deje traducir sin pérdida a un cálculo predicativo de primer o segundo orden, dicha influencia ha sido insignificante. Para confirmar o refutar esta sospecha sería menester, eso sí, poner en claro como *se entienden* a la luz de *Principia Mathematica* las grandes teorías de la matemática clásica (para determinar, luego, el papel que ese particular modo de entenderlas ha desempeñado en el desarrollo ulterior de las mismas). Los tres estudios más notables sobre la obra de Whitehead y Russell —Ramsey 1925, Quine 1941 y Gödel 1944— dejaron inexplorado este tema y no conozco ningún trabajo que lo aborde.

de Black (en Schilpp 1944)— aceptará Russell verla así (Schilpp 1944, p. 691). La versión publicada en el tomo I de *Principia Mathematica* en 1910 es el fruto de una meditación compleja y vacilante, atenta a preocupaciones metafísicas, y ello se le nota. En lo que sigue trazaré a grandes rasgos su desarrollo.[5]

A la teoría cantoriana del transfinito se le imputaban dos paradojas que algunos juzgaban fatales: (i) La paradoja de Cantor: el cardinal del conjunto de todos los conjuntos tiene que ser mayor que cualquier otro cardinal pero, según el Teorema de Cantor, es menor que el cardinal del conjunto de las partes del conjunto de todos los conjuntos. (ii) La paradoja llamada de Burali-Forti: el conjunto de los ordinales está bien ordenado y por lo tanto tiene un ordinal, mayor que cualquier ordinal perteneciente a dicho conjunto, el cual, por ser un ordinal, pertenece sin embargo a ese conjunto. La paradoja de Russell, inspirada por la demostración del Teorema de Cantor (Capítulo 1.6), se deja insertar en este mismo orden de ideas, si entendemos, con Russell (1903, § 68), que las *clases* a que se refiere la paradoja son precisamente los objetos que en la jerga matemática se llaman *conjuntos*. Se distingue, empero, de las dos paradojas citadas en cuanto no envuelve los conceptos de cardinal u ordinal ni ningún otro concepto específicamente matemático, de modo que es propiamente una paradoja de la *lógica universal,* no de una disciplina matemática particular. En su primer intento de resolverla, Russell apela al distingo —introducido en Russell 1903, Cap. VI— entre una clase considerada como pluralidad de objetos (*the class as many*) y una clase considerada como siendo ella misma un objeto (*the class as one*). Damos por descontado —dice Russell— que dondequiera hay una clase-pluralidad también hay una clase-objeto, pero este axioma no tiene que valer universalmente y parecería ser la fuente de la contradicción. “Así pues, con sólo negarlo se superará toda la dificultad” (1903, § 104).[6]

5 Cocchiarella (1980) contiene un análisis más detallado de la compleja evolución de la teoría de los tipos desde 1903 hasta 1910 y explica su curiosa motivación filosófica. Rodríguez Consuegra (1989) ilumina el proceso con la luz que emana de los inéditos de Russell y de paso corrige algunas opiniones de Cocchiarella.

6 El distingo russelliano entre *the class as one* (literalmente, ‘clase como uno’) and *the class as many* (‘clase como muchos’) es afín al distingo entre referencia distributiva y colectiva que utilicé en el Capítulo 1.1. Sea *k* un sustantivo común o frase sustantiva que especifica una clase. Entonces, podemos referirnos *distributivamente* a *cualquier k* (*any k*) o a *cada k* (*every k*) y también podemos referirnos *colectivamente* a *todos los k* (*all k*). Este último modo de referencia englobaría a los *k* en una unidad objetiva, *the*

Como veremos, Russell luego descartará esta solución, debido a la imposibilidad de fijar un criterio racionalmente justificable que determine los casos en que a una clase-pluralidad le corresponde una clase-objeto y los casos en que no le corresponde. Al parecer fue buscando ese criterio que Russell dio con la idea de *tipo lógico*, conservada luego en la solución que adopta finalmente. Para Russell el concepto de clase está estrechamente ligado al concepto de *función proposicional*, que explica así: "$\phi x$ es una función proposicional si, para cada valor de $x$, $\phi x$ es una proposición, bien determinada si $x$ está dado" (1903, § 22). "Una clase puede definirse como todos los términos que satisfacen una cierta función proposicional" (§ 23). En el vocabulario de Russell 1903, un 'término' (*term*) no es una palabra o frase de cierto género, sino "cualquier cosa que pueda ser objeto del pensamiento ... o pueda ser contada como *una*" (§ 47). Así pues, la clase correspondiente a la función proposicional $\phi x$ comprende un objeto $u$ si y sólo si $\phi u$ es una proposición verdadera. Ahora bien, según Russell

> cada función proposicional $\phi x$ ... posee además de su ámbito de verdad (*range of truth*), un ámbito de significación (*range of significance*), esto es, un ámbito dentro del cual $x$ debe hallarse si $\phi x$ ha de ser una proposición, ya sea verdadera o falsa. Los ámbitos de significación forman *tipos*, esto es, si $x$ pertenece al ámbito de significación de $\phi x$, existe una clase de objetos, el *tipo* de $x$, todos los cuales también tienen que pertenecer al ámbito de significación de $\phi x$, como quiera que varíe $\phi$; y el ámbito de significación es siempre, o bien un tipo único, o una suma de varios tipos completos.
>
> (Russell 1903, § 497)

Las seis apretadas páginas que Russell dedica a "La doctrina de los tipos" (1903, Apéndice B) no son un dechado de claridad,[7] pero es bastante claro

---

*class as one*, la clase-objeto. Pero no toda expresión especificadora $k$ tiene esta virtud. Así, si $k$ es la frase sustantiva "clase que no es miembro de sí misma", la paradoja de Russell implica que no puede existir la clase-objeto de *todas* las $k$. Pero —contra lo que Russell va a sostener más tarde— ello no nos impediría hablar con sentido de una $k$, cualquier $k$, cada $k$.

7 Especialmente desconcertante es la relación que establece Russell aquí entre su incipiente doctrina de los tipos y el par de conceptos *clase-como-muchos*/*clase-como-uno*. Según él, una clase-como-uno es un objeto del mismo tipo que los comprendidos en

cómo piensa utilizarla para resolver su paradoja: si el ámbito de significación de una función proposicional $\phi x$ comprende objetos de un tipo dado, entonces su ámbito de verdad —es decir, la clase de los objetos $u$ tales que $\phi u$ es una proposición verdadera— es un objeto de otro tipo diferente. Por lo tanto, si esta clase es designada por la expresión '$\{u\colon \phi u\}$', la expresión '$\phi\{u\colon \phi u\}$' carece de sentido, puesto que $\{u\colon \phi u\}$ no pertenece al ámbito de significación de $\phi x$. Así pues, en el contexto de la doctrina de los tipos, no tiene cabida la noción de que una clase esté o no esté comprendida dentro de sí misma y la paradoja de Russell no se puede siquiera expresar. Las paradojas de Cantor y Burali-Forti también se eliminan por esta vía.

Russell (1906) retorna al tema de las paradojas, comentando un artículo del matemático Hobson. Éste daba por supuesto que un "agregado" (*aggregate*) —hoy diríamos un conjunto— de objetos sólo puede constituirse si hay una *norma* que lo defina, esto es, una condición determinada que cumplan todos sus elementos y sólo sus elementos (Russell aclara que lo que Hobson llama una *norma* es lo mismo que él llama una *función proposicional*).[8] Por eso, Hobson cuestiona el Axioma de Selección adoptado por Zermelo (1904) para demostrar el Teorema del Buen Orden. Según este axioma, si $\mathscr{F}$ es una familia de conjuntos no vacíos, existe un conjunto que con-

---

ella: "Lo que llamamos en el Capítulo VI la clase-como-uno es un individuo, con tal que sus miembros sean individuos: los objetos de la vida diaria, personas, mesas, sillas, manzanas, etc., son clases-como-uno" (1903, § 497). Al tipo superior siguiente pertenece en cambio la respectiva clase-como-muchos. Aparentemente, Russell identifica aquí la clase-como-uno que forman ciertos objetos con lo que llamaríamos hoy la *fusión* o *suma mereológica* de esos objetos (su relación con ellos es como la que tiene una barra de acero con los átomos de hierro y carbono de que consta). En cambio, la clase-como-muchos, en abierto contraste con la caracterización original de este concepto en el Capítulo VI, es tratada aquí como un objeto, de otro tipo que sus miembros, sí, pero *dotado de unidad* y susceptible de agruparse con otros de su mismo tipo para formar nuevas clases (como-uno y como-muchos).

8 Esta identificación me parece demasiado optimista: las *normas* de Hobson son oraciones del inglés o de otro idioma, de las que hay ejemplos por docenas en cualquier libro de matemáticas; en cambio, las *funciones proposicionales* de Russell, que son aplicaciones del reino de los objetos en el reino de las proposiciones, son —como este último— criaturas de su rica inventiva metafísica, cuya misma naturaleza cambia de uno a otro de sus escritos. Más exacto sería decir, pues, que lo que Hobson llama 'norma' es lo que Russell intentaba "reconstruir racionalmente" mediante su concepto de función proposicional.

tiene precisamente un elemento de cada elemento de $\mathscr{F}$, aunque no hay ninguna norma que lo defina. Hobson parece haber pensado asímismo que las paradojas de la teoría de conjuntos nacen de la ausencia o imperfecta especificación de una norma, pero, como Russell hace ver, la situación en este caso es muy diferente:

> La dificultad de los agregados inconsistentes surge de la presencia de una *norma* perfectamente bien definida combinada con la ausencia demostrable del agregado correspondiente. Esto sugiere que una *norma* es una condición necesaria pero no suficiente para la existencia de un agregado; de ser así, la solución completa de [la dificultad de los agregados inconsistentes] estribaría en descubrir las condiciones precisas que una norma debe cumplir para definir un agregado.
>
> (Russell EA, p. 136)

Russell señala que la palabra 'agregado' se usa a veces de un modo que implica la existencia de un orden, y que él utilizará 'clase' (*class*) para referirse a un agregado sin tal implicación. Si toda norma lógicamente bien determinada definiera una clase, entonces existiría la clase $\mathscr{W}$ formada por todos los entes $x$ que cumplen la norma '$x = x$'. Como todo ente cumple esa norma, no puede haber un cardinal mayor que $|\mathscr{W}|$, el cardinal de la clase $\mathscr{W}$. Pero Cantor ha demostrado que, si $\mathscr{W}$ existe, $|\mathscr{W}| < |\mathscr{P}\mathscr{W}|$. Por lo tanto, la norma '$x = x$' no determina una clase. Análogamente, si la norma '$x$ no es una clase que sea un miembro de sí misma' definiera una clase $w$, tendríamos que $w$ es un miembro de sí misma si y sólo si no lo es. Por lo tanto, tampoco esta norma define una clase. De este modo, "aparte de toda opinión sobre la naturaleza de los cardinales y sin hacer consideraciones pertenecientes a la aritmética, podemos probar que hay al menos una norma perfectamente determinada que no define una clase" (EA, p. 139). Russell llama *no-predicativas* (*non-predicative*) a las normas (funciones proposicionales) de una variable que no definen clases, y *predicativas* (*predicative*) a aquéllas que sí las definen; análogamente, una función proposicional de dos variables se llama predicativa si define una relación binaria, no-predicativa si no la define, etc. "Necesitamos, pues, reglas para decidir cuáles normas son predicativas y cuáles no lo son, a menos que adoptemos la opinión (que, como veremos, es muy recomendable) de que *ninguna* norma es predicativa" (p. 141).

Russell bosqueja tres soluciones posibles de este problema, que bautiza (*a*) la teoría del zigzag, (*b*) la teoría de la limitación de tamaño y (*c*) la teoría sin clases (*the no classes theory*). La *teoría del zigzag* supone que si $\phi x$ es una función predicativa, también su negación $\sim\phi x$ es predicativa.[9] Sea $\psi x$ una función no-predicativa. Entonces, no existe una clase $\{x: \psi x\}$. Por lo tanto, cualquiera que sea la función predicativa $\phi x$, el ámbito de verdad de $\psi x$ zigzagueará entre la clase $\{x: \phi x\}$ definida por ella y la clase $\{x: \sim\phi x\}$ definida por su negación. Según Russell, esta es la teoría sugerida en *Principles of Mathematics*, §§ 103 y 104 (EA, p. 146n.). Su desarrollo cabal requiere axiomas que establezcan qué funciones son predicativas. Se parte de la base que todas las funciones proposicionales simples son predicativas y sólo algunas complicadas y extrañas no lo son. Pero, según Russell, cuando intentó precisar esta idea básica tuvo que formular axiomas excesivamente complicados y carentes de toda plausibilidad intrínseca. Además, para seleccionarlos no tuvo otra guía que la necesidad de evitar contradicciones, la cual "por sí sola, es un principio muy insuficiente, pues nos deja expuestos siempre al riesgo de que nuevas inferencias generen contradicciones" (EA, p. 147).

Russell atribuye la *teoría de la limitación de tamaño* a su amigo el matemático Philip Jourdain. Según ella una función proposicional define una clase no cuando su fórmula es suficientemente sencilla, sino cuando efectivamente *deslinda* una colección de objetos, y no implica —como las funciones '*x* es un cardinal' y '*x* es un ordinal'— que todo límite que se designe será rebasado. Desde este punto de vista, obviamente, si $\phi x$ es predicativa, $\sim\phi x$ no puede serlo. Russell opina que "una gran dificultad de esta teoría consiste en que no nos dice hasta dónde es legítimo avanzar en la serie de los ordinales. Pudiera ser que ya $\omega$ sea ilegítimo, en cuyo caso todas las clases propiamente tales (*all proper classes*) serían finitas" (EA, p. 153).[10]

9 Uso aquí la notación de Russell 1903. Russell 1906 designa una función proposicional cualquiera, predicativa o no, con una expresión formada por una minúscula griega seguida de un signo de admiración y una minúscula latina cursiva, vgr. $\phi!x$. Pero en *Principia Mathematica* el signo de admiración a la derecha de una minúscula griega distingue a cierto género de funciones (llamadas "predicativas", pero en una nueva acepción que explico en la nota 20). Para no turbar al lector con tantas notaciones incompatibles, me abstengo de emplear la de Russell 1906.

10 Obsérvese que, como Russell equipara sus *clases* a lo que los matemáticos llaman *conjuntos*, tiene que llamar *clases propias* a lo que los matemáticos —desde von

En 1906, todas las simpatías de Russell van a la *teoría sin clases*. El escrito que comentamos concluye con una nota suplementaria, fechada el 6 de febrero de 1906, en la que dice que sus últimas investigaciones lo han persuadido de que la teoría sin clases proporciona "la completa solución" de las dificultades suscitadas por las paradojas (EA, p. 164). Se refiere, sin duda, a los resultados que presenta en su artículo "Sobre la teoría sustitucional de las clases y las relaciones" (1906a), sometido a la London Mathematical Society el 24 de abril de 1906, pero retirado antes de que saliera impreso, porque dejó de satisfacerlo (se publicó póstumamente en 1973). El nombre 'teoría sustitucional' alude al método de sustitución de expresiones con que propone eliminar —o hacer ontológicamente inocuos— los ingredientes del lenguaje que aparentemente hacen referencia a las clases.[11] Dicho método le fue sugerido seguramente por su brillante análisis de las descripciones definidas (Russell 1905), que paso a explicar.

Afligido por la dificultad de asignar una denotación a expresiones tales como 'la esposa favorita del obispo de Roma' o 'el río de oro derretido que desemboca en el Orinoco', que describen algo que no existe, Russell optó por concebirlas como abreviaturas que sirven para representar en contextos de un cierto género a otras expresiones más largas, las cuales, a su vez, no denotan nada. Concretamente, cuando se dice que

(1) El río de oro derretido que desemboca en el Orinoco es más caudaloso que el Guadalquivir

lo que se expresa, según Russell, es simplemente que

---

Neumann— llaman *clases impropias* (*propias* son precisamente las clases que *no son* conjuntos). Recordemos de paso que —como se vio en la Sección 1.8.4— la teoría de conjuntos de von Neumann da una respuesta precisa a la dificultad mencionada por Russell: una clase es "demasiado grande" —y por lo tanto es *propia* en el sentido de von Neumann, *impropia* en el sentido de Russell— si y sólo si es equinumerosa con la clase de todos los conjuntos.

11 Como Russell abandona muy pronto la teoría de 1906a pero sigue simpatizando con la idea de la eliminación de las clases, los expertos distinguen entre la 'teoría sin clases' y la 'teoría sustitucional de las clases'. Pero, comparando textos, no hay duda de que la 'teoría sin clases' bosquejada en Russell 1906 (EA, pp. 154-56) y aludida en la nota suplementaria al final (EA, p. 164), no es otra que la teoría sustitucional explicada en 1906a.

(1*) Existe un objeto $x$ tal que (i) $x$ es un río de oro derretido que desemboca en el Orinoco, (ii) $x$ es más caudaloso que el Guadalquivir y (iii) si cualquier objeto $y$ es un río de oro derretido que desemboca en el Orinoco, entonces $y$ es idéntico a $x$.

En (1*), la descripción del río de oro figura sólo como predicado, no como frase nominal. La oración es falsa porque *no hay* nada que satisfaga ese predicado, y no, como sugiere (1), porque el objeto denotado por la frase nominal sea *un río no más caudaloso que el Guadalquivir*.

No es raro que Russell, después de eliminar con tanta soltura la denotación de las descripciones definidas, se sintiera animado a hacer lo propio con los nombres de clases. Se trata "de proveer un modo de interpretar las aseveraciones corrientes sobre clases sin suponer que las clases son entes" (1906c, p. 200). Por esta vía, resultará que todas las proposiciones significativas en que se mencionan clases pueden entenderse como proposiciones acerca de sus miembros, esto es, sobre algunos o todos los objetos que satisfacen cierta función proposicional $\phi x$. Sólo las proposiciones que no se dejan entender así dan lugar a contradicciones. "Por lo tanto, es natural suponer que las clases son meramente abreviaturas lingüísticas o simbólicas" (Ibid.). Russell reconoce que el procedimiento de eliminación propuesto por él es complicado, un truco técnico que debe reemplazarse por otro más conveniente. Las siguientes indicaciones, basadas en su primera presentación sumaria de la teoría (Russell 1906, en EA, pp. 154-56), darán una idea de lo que se trata.

Sea $p$ una proposición y $p(x/a)$ lo que se obtiene cuando $x$ reemplaza a $a$ en todos los lugares en que $a$ figura en $p$; $p(x/a)$ nos da, para distintos valores de $x$ lo que solíamos llamar los distintos valores de una función proposicional.[12] Si $b$ es un ente cualquiera que no es un ingrediente de $p$ y $q = p(b/a)$, tenemos que la aseveración '$q(x/b)$ es verdadera para todo valor de $x$'

[12] Lo anterior traduce literalmente palabras de Russell. No pretendo que sean claras. Como una *proposición* es un estado de cosas —real o posible— la letra $a$ designa aquí un ingrediente objetivo del mismo. La variable $x$, en cambio, no puede sino ser un objeto lingüístico: un pronombre si está ligada, la indicación de un "hueco" (en el sentido de Frege) si está libre ¿Cómo se puede reemplazar un ingrediente objetivo de un estado de cosas con un "hueco"? ¿Cuál es el modo de ser del producto de tal procedimiento? Incapaz de responder a estas preguntas, las tomo como un indicio más de la confusión que, en diversas modalidades, acompañó a Russell durante casi toda su carrera filosófica.

equivale a '$p(x/a)$ es verdadera para todo valor de $x$'. Cabe, pues, decir que la aseveración '$p(x/a)$ es verdadera para todo valor de $x$' no depende del sujeto $a$, sino "sólo de la *forma* de $p$" (EA, p. 155). Russell propone que usemos aseveraciones de este tipo en vez de mencionar funciones proposicionales o clases determinadas por ellas.

> Por ejemplo, en vez de '$\phi$ es una función unitaria' (esto es, 'Hay un y sólo un $x$ tal que $\phi x$ es verdad'), tendremos 'Hay un ente $b$ tal que $p(x/a)$ es verdadera si y sólo si $x$ es idéntico a $b$'. No existirá ya, pues, un ente como el número 1, aislado; pero podremos definir lo que significamos al decir 'Una y sólo una proposición del tipo $p(x/a)$ es verdadera (dados $p$ y $a$)'. En vez de decir 'La clase $u$ es una clase que tiene sólo un miembro', diremos [ . . . ] 'Hay un ente $b$ tal que $p(x/a)$ es verdadera si y sólo si $x$ es idéntico a $b$'. Aquí los valores de $x$ para los cuales $p(x/a)$ es verdadera reemplazan a la clase $u$; pero no presuponemos que estos valores forman colectivamente un ente único que es la clase compuesta de ellos.
>
> (Russell EA, p. 155)[13]

Russell observa que no es difícil reemplazar la antigua escritura por la nueva, pero que la demostración de teoremas de existencia se ve entonces muy entorpecida. En 1906 esperaba "elaborar esta teoría hasta el punto en que se vea claramente cuánto preserva de las matemáticas y cuánto nos fuerza a abandonar" (EA, p. 156). Como 'clases' y 'conjuntos' ('agregados') son para Russell lo mismo, la negación de la existencia de tales entes imprime a su filosofía de las matemáticas un rumbo muy distinto del que tomó por esos mismos años la filosofía de los matemáticos. Como vimos en el Capítulo 1.7, Zermelo (1908a) postula la existencia de un mínimo de conjuntos que le parecían imprescindibles para hacer matemáticas, y presume que su teoría es inocente de contradicciones mientras no se la pruebe culpable.[14] La co-

13 Como advertí en la nota 9, altero un poco la notación de Russell: él escribe $\phi!x$ donde yo $\phi x$. También Lackey, el editor de EA, se ha tomado libertades con la notación original: Russell ponía el "numerador" sobre el "denominador" en la "fracción" $x/a$, con lo cual podía prescindir de los paréntesis que Lackey y yo utilizamos para separarla de la $p$. Curiosamente, después de reformar esta notación de Russell en el texto de 1906, Lackey se ciñe fielmente a ella en el texto de 1906c (el original inglés de 1906b).

14 Subrayo que los axiomas de Zermelo no se eligen, como en la teoría russelliana del zigzag, sólo con vistas a prevenir las contradicciones conocidas. Zermelo tiene un cometido —hacer matemáticas— y postula lo que necesita para eso. Su selección se ha

rriente central de la matemática conjuntista ha seguido su ejemplo con una temeridad que aterra a los temperamentos metafísicos, ansiosos de cobijarse en una verdad totalitaria e inquebrantable.

Whitehead y Russell declaran reiteradamente que el sistema de *Principia Mathematica* es neutral con respecto a las clases: no presupone su existencia ni su inexistencia (PM, I, 24, 187). Pero la escritura conceptual utilizada en el libro no cuenta con los medios para nombrar tales objetos; una expresión como $\hat{x}(\phi x)$—que se lee "la clase de los objetos $x$ que satisfacen la condición $\phi$"— *simula* hacerlo, pero es un símbolo incompleto, definido siguiendo un patrón análogo al arriba descrito (véase la definición contextual de '$\hat{x}(\phi x)$' en la nota 35). Es verdad que en *Principia Mathematica* la eliminación de las clases no lleva la voz cantante en la solución de las paradojas;[15] pero opera tras bastidores. La solución se funda expresamente en el "principio del círculo vicioso", que Russell (1906b, 1908) toma de Poincaré. Pero dicho principio es insostenible si de veras hay clases (véase la nota 41 y el texto que remite a ella).

Henri Poincaré, matemático profundo y agudo filósofo, no simpatizaba con el "cantorismo" (como llamaba a la teoría de conjuntos) y menos aún con la "logística" (esto es, la nueva lógica premunida de escritura conceptual y ávida de anexarse las matemáticas). El fenómeno de las paradojas le parece una señal clarísima de que cantoristas y logísticos andan perdidos. En el segundo de tres artículos sobre "Las matemáticas y la lógica" (1905/1906),[16] Poincaré introduce el tema con marcado sarcasmo:

---

probado duradera. En cambio, Russell, que buscaba *certificar* —como si hiciera falta— las matemáticas hechas por otros, daba solamente con axiomas implausibles, inspirados por un principio que él mismo juzgaba insuficiente (1906, en EA, p. 147, citado arriba en la p. 183).

15 Esto puede deberse a que Whitehead no favorecía la eliminación de las clases. El 22 de febrero de 1906 —o sea, dos semanas después de la fecha de la nota suplementaria en que Russell abraza sin reservas esa alternativa— Whitehead le escribió protestando contra su teoría sustitucional, que "funda toda la matemática en un artificio tipográfico (*a typographical device*) y de este modo contradice las doctrinas principales del tomo I [= Russell 1903]" (citado por Lackey en Russell, EA, p. 131; cf. las otras citas y comentarios de Lackey en las pp. 131-32).

16 Una versión revisada de estos ensayos forma los Capítulos III, IV y V del libro II de *Science et méthode* (Poincaré 1908). A ella remiten mis referencias.

> ¿Pueden las matemáticas ser reducidas a la lógica sin recurrir a principios que le sean propios? Existe una escuela llena de ardor y de fe que se esfuerza por establecerlo. Tiene un lenguaje especial en el que las palabras se sustituyen por signos. Este lenguaje no lo comprenden más que algunos iniciados, de manera que los profanos están dispuestos a inclinarse ante las decisivas afirmaciones de los adeptos. No me parece inútil examinar estas afirmaciones más de cerca a fin de ver si justifican el tono perentorio con que se las presenta.
>
> (Poincaré, CM, p. 111)[17]

Poincaré reúne bajo el nombre de *antinomias*[18] *cantorianas* las paradojas que hemos mencionado ya, en las que figura decisivamente la noción de conjunto o la de clase, con otras que aún no hemos mencionado, en que estas nociones no desempeñan ningún papel. Tiende un puente entre ambos grupos la paradoja de Richard (1905), en la cual figura la noción de conjunto, pero de modo inesencial. Hela aquí: sea $E$ el conjunto de los números reales del intervalo (0,1) que se pueden caracterizar en nuestro idioma mediante un número finito de palabras. $E$ es evidentemente un conjunto numerable. Sea $f\colon \mathbb{N} \to E$ una enumeración de $E$. Designemos con $\alpha_{hk}$ al $k$-ésimo dígito de la expansión decimal infinita de $f(h) \in E$. Caracterizaré ahora un número real $\beta$ del intervalo (0,1) mediante la condición siguiente: Sea $\beta_n$ el $n$-ésimo dígito de la expansión decimal infinita de $\beta$; entonces $\beta_n = 5$ si $\alpha_{nn} \neq 5$ y $\beta_n = 6$ si $\alpha_{nn} = 5$. Evidentemente, $\beta \notin E$, puesto que, para cada $n \in \mathbb{N}$, $\beta$ difiere de $f(n)$ en el $n$-ésimo dígito de la respectiva expansión decimal infinita. Más, por otra parte, es claro que $\beta \in E$, puesto que acabamos de caracterizarlo en castellano mediante un número finito de palabras. El argumento de Richard

17 Cf. también este otro pasaje, referente a la paradoja de Burali-Forti: "Mientras se trate de demostrar que uno es un número, la pasigrafía basta; pero si se presenta una dificultad, si hay una antinomia que resolver, la pasigrafía se torna impotente" (Poincaré, CM, p. 123). 'Pasigrafía' —esto es, 'escritura para todos' (o 'para todo')— es el nombre que daba Peano a su escritura conceptual.

18 'Antinomia' —es decir, antilegalidad— llamó Kant al "conflicto de la razón pura consigo misma" que estudia en su *Crítica de la razón pura*. Dicho conflicto se manifiesta, según Kant, en la demostración de varios pares de aseveraciones contradictorias. Conforme a un precedente establecido por el propio Kant, la palabra 'antinomia' se usa comúnmente para designar a cada uno de estos pares. Por ejemplo, la segunda antinomia kantiana consta de la tesis, "Todo cuerpo consta de partes indivisibles" y la antítesis, "Todo cuerpo es indefinidamente divisible".

apela, es cierto, a la noción de conjunto y se vale una vez más del método diagonal de Cantor para generar una contradicción, pero se puede formular una paradoja esencialmente análoga a ésta sin esos recursos. Poincaré (CM, p. 144) cita la siguiente paradoja que atribuye Russell, pero que según éste le fue sugerida por G. G. Berry (Russell LK, p. 60n.). Considérese *el entero positivo más pequeño que no se puede describir en castellano con menos de diechiocho palabras*. Como, obviamente, hay enteros positivos que sólo pueden describirse con 18 palabras o más, tiene que haber uno entre ellos —llamémosle $\alpha$— que sea menor que todos los demás. Según esto, $\alpha$ no puede describirse con menos de 18 palabras. Sin embargo, la frase en cursiva lo describe inequívocamente, y tiene sólo 17 palabras. Como se puede ver, la contradicción no incide en una relación como la simbolizada con '$\in$', que podemos pensar que subsiste en virtud de la propia naturaleza de las cosas relacionadas, sino en la relación obviamente convencional entre un término y la expresión que lo define. Este giro lingüístico se advertía ya en la primera formulación de la paradoja de Russell en la carta que éste escribió a Frege el 16 de junio de 1902: "Sea *w* el predicado 'es un predicado que no puede predicarse de sí mismo' ¿Se puede acaso predicar *w* de sí mismo? De cada respuesta se infiere lo contrario" (Frege, WB, p. 211; *vide supra*, p. 143, nota 22).[19]

Según Poincaré, Richard (1905) ofrece la verdadera solución de las "antinomias cantorianas". Consideremos una vez más el conjunto *E*, men-

19 Al mismo género pertenece la siguiente paradoja publicada en Russell 1903, Apéndice B, § 500: Sea *m* una clase de proposiciones; la proposición 'todas las proposiciones de la clase *m* son verdaderas' puede o no pertenecer a la clase *m*; sea *w* la clase de todas las proposiciones de la forma indicada que *no* pertenecen a la clase que mencionan y sea *p* la proposición 'todas las proposiciones de la clase *w* son verdaderas'; entonces *p* pertenece a *w* si y sólo si no pertenece a *w*. Russell señala allí que esta paradoja no puede resolverse con la teoría de los tipos (simple, no ramificada) propuesta en dicho Apéndice B.

El giro lingüístico se consuma en la siguiente paradoja, descubierta por Kurt Grelling: digamos que una palabra es *autológica* si tiene la propiedad que ella expresa o pertenece a la clase que ella nombra, *heterológica* en caso contrario. Por ejemplo, 'breve', 'pentasílaba', 'esdrújula' y 'sustantivo' son palabras autológicas, 'larga', 'tetrasílaba', 'aguda' y 'adverbio' son heterológicas. Ahora bien, la palabra 'heterológica' ¿es heterológica o autológica? Evidentemente, si es heterológica es autológica y si es autológica es heterológica (Grelling y Nelson 1908).

cionado en la presentación de la paradoja debida a este autor. *E* es el conjunto de los números reales del intervalo (0,1) que se pueden caracterizar en nuestro idioma mediante un número finito de palabras *sin hacer referencia al conjunto E*. La cláusula en cursiva tiene que subentenderse en la definición de *E* pues de otro modo dicha definición contendría un círculo vicioso. "No se puede definir *E* por el conjunto *E* mismo" (Poincaré, CM, p. 146). Pero entonces no hay duda que $\beta \notin E$ y que es falsa la conclusión antinómica $\beta \in E$. En efecto, aunque el número β se ha definido con un número finito de palabras, esas palabras contienen una referencia a *E*. Según Poincaré, "la misma explicación vale para las otras antinomias, como es fácil verificar" (pp. 146ss.). De este modo, habríamos por fin hallado el criterio —que Russell (1906) decía haber buscado en vano— para distinguir las definiciones (normas, funciones proposicionales) predicativas de las que no lo son: "*Las definiciones que deben ser consideradas como no predicativas son las que contienen un círculo vicioso*" (Poincaré, CM, p. 147). Aquí 'predicativa' designa —siguiendo a Russell (1906)— las definiciones que determinan una clase. Sin embargo, el texto recién citado de Poincaré —en cursiva también en el original— se entenderá más tarde como una *definición* de los términos técnicos 'predicativo' y 'no predicativo' o 'impredicativo', como se suele decir. Lo que en lógica y matemáticas se llama hoy *impredicatividad* —y que algunos autores consideran como un abuso del lenguaje o al menos del pensamiento— consiste en esa forma de circularidad que Poincaré reputó viciosa y que Russell intentará caracterizar con más precisión.[20]

20 Véase el Capítulo 1.7, nota 4. Para mayor confusión, Russell (1908) emplea la palabra 'predicativo' en una tercera acepción. Como, en virtud de la teoría sin clases, no hay funciones proposicionales predicativas en el sentido de Russell (1906), y en virtud de la sintaxis adoptada, no es posible representar en la escritura conceptual una función proposicional impredicativa en el sentido de Poincaré (1906), la palabra ha quedado vacante y se la redefine entonces como sigue: Una función de orden *n* y una sola variable es *predicativa*, si esa variable es de orden $n - 1$; una función de varias variables es *predicativa*, si al reemplazar por constantes todas sus variables excepto una se obtiene una función predicativa de la variable restante (Russell LK, p. 78). Más adelante veremos qué es lo que aquí se llama el orden de una función o de una variable. Whitehead y Russell 1910/13 reiteran esta definición (PM, I, 53), pero dan también otra más sencilla que, en virtud de la sintaxis adoptada en ese libro, resulta ser equivalente a la anterior: "Se dice que una función es *predicativa* cuando es una matriz", esto es, cuando "no envuelve variables ligadas (*apparent variables*)" (PM, I, 164, 163).

En "La lógica matemática basada en la teoría de los tipos" (1908), Russell ofrece una solución detallada de las paradojas. Muchos pasajes de este artículo se reproducen sin variación en *Principia Mathematica* y aunque es posible señalar diferencias de doctrina entre ambas obras, aquí no les prestaremos atención.[21] El artículo empieza con una lista de siete paradojas, encabezada por la clásica *paradoja del mentiroso*, que no habíamos encontrado en la literatura comentada hasta aquí, pero que ilustra con especial elocuencia la circularidad denunciada por Poincaré: si digo "estoy mintiendo" lo que digo es falso si es verdad y es verdad si es falso.[22] Según Russell, cada una de las paradojas de su lista da por supuesta "una totalidad que, si fuera legítima, instantáneamente se incrementaría con nuevos miembros definidos en términos de ella misma" (LK, p. 63). Esta observación debe llevarnos, según él, a adoptar la regla siguente:

> 'Lo que envuelve el *todo* de una colección (**all** *of a collection*) no puede ser un miembro de esa colección'; o, recíprocamente: 'Si, en caso que cierta colección tuviera un total, tendría miembros definibles sólo en términos de ese total, dicha colección no tiene un total'.
>
> (Russell LK, p. 63)

21 Rodríguez Consuegra 1989, pp. 153ss. estudia el significado de esas diferencias a la luz de manuscritos inéditos que Russell redactó por esos años. Russell 1910 es una exposición de la teoría de los tipos contemporánea de PM, destinada a rebatir las críticas de Poincaré a Russell 1908. Este artículo se publicó en francés; el original inglés se hallará en Russell, EA, pp. 215-52.

22 Cf. Cicerón, *Academica*, II.95: "Si te mentiri dicis idque verum dicis, mentiris?" Atribuida por Diógenes Laercio (II.108) a Eubúlides de Mileto, un contemporáneo de Aristóteles, la paradoja del mentiroso debe su difusión sobre todo a la epístola de San Pablo a Tito (i.12; cf. Calímaco, H. 1.7), donde se alude a cierto "profeta", quien habría dicho que "los cretenses son siempre mentirosos" (Κρῆτες ἀεὶ ψεῦσται), siendo así que él mismo era cretense. En el artículo "Insolubilia" del *Dictionary of Philosophy and Psychology* de Baldwin (1901-1905), C. S. Peirce presenta una versión de la paradoja como paradigma de "una clase de sofismas en los cuales se plantea una cuestión de tal índole que, ya sea que se responda afirmativa o negativamente a ella, un argumento formalmente impecable demostrará que la respuesta es falsa" (CP, 2.618). Peirce menciona dos métodos de solución propuestos en la Edad Media. Uno, adoptado por Ockam, arguye que "ninguna proposición puede aseverar nada acerca de sí misma". En "Les paradoxes de la logique" (1906b), Russell cita el artículo de Peirce a propósito justamente de este método de Ockam (EA, p. 196, nn. 2 y 3). Obsérvese que el original inglés de "Les paradoxes de la logique" sigue a Peirce también en su título: "On 'Insolubilia' and their Solution by Symbolic Logic" (1906c).

*Principia Mathematica* repite textualmente los dos enunciados de esta regla y le da un nombre: "*el principio del círculo vicioso*" (PM, I, 37). Este principio impone un orden jerárquico a las funciones proposicionales. Según Whitehead y Russell, la característica esencial de una función es su *ambigüedad*. Cuando hablamos de '$\phi x$' sin especificar $x$, nos referimos a un valor indeterminado de la función. Si los valores de la función son $\phi a$, $\phi b$, $\phi c$, ..., podemos decir que '$\phi x$' *denota ambiguamente* a $\phi a$, $\phi b$, $\phi c$, ...

> Por lo tanto, "$\phi x$" sólo tiene un significado bien definido (es decir, bien definido excepto en cuanto es de su esencia ser ambigua) si los objetos $\phi a$, $\phi b$, $\phi c$, etc., están bien definidos. En otras palabras, una función no es una función bien definida a menos que todos sus valores ya estén bien definidos. De esto se desprende que ninguna función puede tener entre sus valores algo que presuponga la función, pues, si lo tuviera, no podríamos considerar que los objetos ambiguamente denotados por la función están definidos mientras la función no estuviera definida, mientras que, a la inversa, como acabamos de ver, la función no puede estar definida mientras no lo estén sus valores. Este es un caso particular, pero tal vez el más fundamental, del principio del círculo vicioso.
>
> (Whitehead y Russell, PM, I, 39)

Imitando a Whitehead y Russell, designaré a la función cuyos valores son $\phi a$, $\phi b$, $\phi c$, etc. con el símbolo '$\phi\hat{z}$'. A la luz del texto citado, es claro que $\phi\phi\hat{z}$ no puede ser un valor de esa función. Más aún, una expresión como '$\phi\phi\hat{z}$' simplemente no tiene sentido.[23] Así, la paradoja russelliana del predicado ' ... es un predicado que no es predicable de sí mismo' no puede siquiera enunciarse. En efecto, si usamos la letra '$\psi$' para simbolizar dicho predicado, su definición sería: $\forall\phi\hat{x}(\psi\phi\hat{x} \leftrightarrow \sim\phi\phi\hat{x})$. Pero la fórmula a la derecha del signo de equivalencia es inadmisible.

Russell (1908) resucita la terminología que ya le vimos utilizar en 1903: los objetos que forman el *ámbito de significación* de una función proposicional $\phi\hat{z}$ —esto es, los objetos $a$, $b$, $c$, ... tales que $\phi a$, $\phi b$, $\phi c$, ... son valores de $\phi\hat{z}$— constituyen un *tipo* de objeto al cual $\phi\hat{z}$ misma no pertenece. La función $\phi\hat{z}$, por su parte, forma con otras funciones un tipo diferente que es el ámbito de significación de funciones proposicionales de un tercer tipo, y así

[23] Dicho sea de paso, en el libro de Whitehead y Russell el resultado de reemplazar $x$ por $\phi\hat{z}$ en $\phi x$ no se escribe $\phi\phi\hat{z}$, sino $\phi(\phi\hat{z})$, una inconsecuencia que prefiero no imitar.

sucesivamente. Whitehead y Russell (PM, I, 47s.) aducen dos razones —nada concluyentes a mi modo de ver— para probar que tales tipos no pueden tener un elemento en común.[24] Por lo tanto, si $\phi a$ es un valor de la función $\phi\hat{z}$ y $\Psi\phi\hat{z}$ es un valor de una función $\Psi\hat{y}$, $\Psi a$ no puede ser un valor de $\Psi\hat{y}$ y la expresión '$\Psi a$' no tiene sentido. Llamaré 'individuos' —como Whitehead y Russell— a los objetos que no son funciones proposicionales ni proposiciones.[25] Según nuestros autores, todos los individuos forman un solo tipo, que

[24] La primera razón propuesta es esta: Una función no puede servir de argumento a otra cuyo ámbito de significación incluya objetos individuales, porque una función no es un objeto definido sino "una mera ambigüedad a la espera de una determinación [ . . . ], y obviamente no la obtiene con sólo reemplazar a algo determinado en una proposición". Por eso los nombres 'Whitehead' y 'Poincaré' pueden sustituir a la variable *x* en la función proposicional '*x* es un matemático inglés', generando, respectivamente, una proposición verdadera y una falsa, pero si reemplazamos la *x* por el nombre de una función proposicional como, por ejemplo, '*z* es una estrella de quinta magnitud' el resultado carece de sentido. Whitehead y Russell reconocen que el citado argumento no se aplica a una función proposicional de una o más variables si todas estas están ligadas (de modo que lo que tenemos entre manos no es propiamente una función proposicional sino una proposición generalizada). Sirva '$(x).Px \supset Kx$' para abreviar la proposición generalizada 'Todo planeta describe una elipse en uno de cuyos focos está el sol'. Entonces, obviamente, la expresión '$(x).Px \supset Kx$ es un matemático inglés' carece de sentido, aunque las partes de que consta no adolecen de indefinición. "Necesitamos, entonces, una nueva objeción, a saber, la siguiente: Una proposición no es un ente singular, sino una relación entre varios; por lo tanto, *una aseveración en que figure una proposición como sujeto sólo será significativa si puede reducirse a una aseveración sobre los términos que figuran en la proposición*. [ . . . ] Pero esto no es posible en el caso de una aseveración tal como '*p* es un hombre', donde *p* es una proposición. Por eso '$\{(x).\phi x\}$ es un hombre' carece de sentido" (PM, I, 48; cursiva mía). El argumento depende de la oración que he destacado en cursiva. No logro imaginarme qué pudo inducir a los autores a pensar que esa oración es verdadera. El ejemplo siguiente demuestra que no lo es: 'El Teorema de Pitágoras se deduce de los axiomas de Hilbert para la geometría euclidiana'. La relación de deducibilidad que aquí se afirma que subsiste entre una cierta proposición y un determinado grupo de proposiciones no puede aseverarse de los términos —puntos, rectas, etc.— que figuran en dichas proposiciones, ni siquiera es equivalente a una relación que subsista entre esos términos.

[25] PM, I, 51, 132. En la p. 162 los autores dicen: "Podemos explicar un individuo como algo que existe por su propia cuenta; entonces, obviamente no es una proposición, puesto que las proposiciones, según se explicó en el Capítulo II de la Introducción (p. 43), son símbolos incompletos, que no tienen significado sino cuando de las usa [en un contexto]." Pero en la p. 161 advierten que "en la práctica no es necesario saber qué objetos pertenecen al tipo más bajo, ni si el tipo más bajo de variable que figura en un

llamaré '0'. Si $\phi\hat{z}$ es una función proposicional de una variable cuyo ámbito de significación es el tipo $t$, digo que $\phi\hat{z}$ es del tipo ($t$). A la luz de lo dicho, parecería que los tipos de las funciones de una variable —los atributos— forman una jerarquía isomórfica a la serie de los números naturales: (0), ((0)), (((0))), . . . En esta jerarquía, cada tipo sería el ámbito de significación de las funciones del tipo siguiente. En cambio, las funciones proposicionales de dos o más variables —las relaciones— no podrían en ningún caso ordenarse con tanta nitidez. Sea $\phi\hat{z}_1 \dots \hat{z}_n$ una función de $n$ variables y $\phi a_1 \dots a_n$ una proposición obtenida desambiguando esa función. Sean $t_1, \dots, t_n$ los tipos —posiblemente diversos— a que pertenecen respectivamente los objetos denotados por los términos $a_1, \dots, a_n$. El ámbito de significación de $\phi\hat{z}_1 \dots \hat{z}_n$ sería entonces el producto cartesiano $t_1 \times \dots \times t_n$. Digamos entonces que $\phi\hat{z}_1 \dots \hat{z}_n$ es una función del tipo $(t_1, \dots, t_n)$. Por ejemplo, si $\psi ab$ se lee 'Pedro ama a Teresa', $\psi\hat{z}\hat{x}$ es una relación binaria del tipo (0,0). Si $\phi acd$ dice que 'Pedro antepone el patriotismo al amor', tenemos que $a$ es un objeto del tipo 0, $c$ es un atributo de individuos, esto es, un objeto del tipo (0), y $d$ es justamente la relación del tipo (0,0) que antes llamé $\psi\hat{z}\hat{x}$. Por lo tanto, $\phi\hat{x}\hat{y}\hat{z}$ es una función del tipo (0,(0),(0,0)). Como es concebible que haya, por ejemplo, funciones del tipo

$$(((0),(0)),(0,0,(0,(0))),((((0))),0))$$

y de otros aún más endiablados, no podemos ordenarlas numéricamente de un modo natural.

El sistema descrito en el párrafo anterior se llama —¡increíble pero cierto!— la *teoría simple de los tipos*. Como ya indiqué, ella disuelve sin dificultad la paradoja de Russell (también las de Cantor y Burali-Forti). Pero, como veremos enseguida, no es compatible con el principio del círculo vicioso, por lo cual Whitehead y Russell, en la primera edición de *Principia Mathe-*

---

dado contexto es el de los individuos o es otro. Pues en la práctica sólo importan los tipos *relativos* de las variables; así el tipo más bajo que figura en un dado contexto puede llamarse 'de los individuos' por lo que hace a ese contexto. [ . . . ] Lo esencial es el modo cómo los otros tipos se generan a partir de individuos, como quiera que esté constituido el tipo de los individuos."

*matica*, adoptan la llamada *teoría ramificada de los tipos*. La cita siguiente explica por qué no pueden aceptar la idea de que las funciones que comparten un dado ámbito de significación formen entre todas un *tipo*:

> Sea $f(\phi\hat{z},x)$ una función de las dos variables $\phi\hat{z}$ y $x$. Entonces si, manteniendo a $x$ fija por el momento, aseveramos esto con todos los valores posibles de $\phi$, obtenemos una proposición:
>
> $$(\phi).f(\phi\hat{z},x).$$
>
> Aquí, si $x$ es una variable, tenemos una función de $x$; pero como esta función envuelve una totalidad de valores de $\phi\hat{z}$, no puede ser ella misma uno de los valores incluidos en esa totalidad, en virtud del principio del círculo vicioso. Se sigue que la totalidad de valores de $\phi\hat{z}$ a que concierne $(\phi).f(\phi\hat{z},x)$ no es la totalidad de todas las funciones en que $x$ puede figurar como argumento, y que no existe una totalidad de todas las funciones en que $x$ puede figurar como argumento.
>
> (Whitehead y Russell, PM, I, 48–49)[26]

Sea $t$ un tipo cualquiera y $\mathscr{F}$ una colección de funciones cuyo ámbito de significación es $t$. Si $a$ es un objeto cualquiera del tipo $t$, entonces la oración '$a$ satisface todas las funciones $\phi\hat{z}$ de la colección $\mathscr{F}$' expresa una proposición. Si en ella reemplazamos la constante $a$ por la variable $x$ obtenemos una fórmula representativa de una función proposicional que llamaré $\psi\hat{z}$. Pero entonces, aunque el ámbito de significación de $\psi\hat{z}$ también es el mismo tipo $t$, el principio del círculo vicioso no permite que $\psi\hat{z}$ pertenezca a $\mathscr{F}$, puesto que $\psi\hat{z}$ se refiere a la totalidad de $\mathscr{F}$.[27] No hay derecho a hablar de *un* tipo

26 Pido al lector disculpas por el simbolismo confuso e inconsecuente, pero no me ha parecido lícito alterarlo en esta cita textual. Nótese que la variable $\phi\hat{z}$ de la primera oración pasa a ser $\phi$ en la oración siguiente; luego, en la fórmula destacada, reviste ambas formas, tocándole al lector adivinar que $\phi\hat{z}$ es la variable ligada por el cuantificador universal $(\phi)$. (Cien páginas más adelante, en PM, I, 165, Whitehead y Russell nos explican que adoptaron esta práctica "en aras de la brevedad"; aunque ella responde también a un motivo menos inocente, como se verá en la nota 35). Tampoco es justo, después que se ha dicho que $x$ es una *variable*, hablar de "las funciones en que $x$ puede figurar como argumento": un *argumento* de una *función* es en cada caso un *objeto determinado* o, metonímicamente, un término *constante* que lo designe.

27 Si una función $\phi\hat{y}$ pertenece a $\mathscr{F}$ si y sólo si $f\phi\hat{y}$, la fórmula representativa de la función que he llamado $\psi\hat{z}$ debe escribirse $(\phi\hat{y})(f\phi\hat{y} \supset \phi x)$ en la notación de PM.

(0) al que pertenecerían *todas* las funciones cuyo ámbito de significación es el tipo 0 de los individuos. *Cualquier* totalidad de funciones de una variable que admitan individuos como argumento tiene necesariamente que *excluir* algunas funciones con ese ámbito de significación. En general, para satisfacer el principio del círculo vicioso, las funciones que tienen un mismo ámbito de significación tienen que repartirse en diversos "órdenes", ninguno de los cuales incluye una función que se refiera a todo ese orden.[28]

Whitehead y Russell no dan una caracterización global precisa de su sistema "ramificado". En vez de intentarla por mi cuenta, me limitaré a resumir las indicaciones que ellos juzgaron suficientes para darse a entender.[29]

28 La palabra 'orden' (*order*) se emplea de este modo en PM, *12. En el pasaje de la Introducción arriba parafraseado, Whitehead y Russell escriben 'tipo' (*type*). Hablan de lo que llaman '*a*-functions', esto es, las funciones proposicionales cuyo ámbito de significación comprende un dado objeto *a*. El texto original dice así: "What is necessary therefore in order to avoid vicious-circle fallacies, is to divide our *a*-functions into 'types,' each one of which contains no functions which refer to the whole of that type" (PM, I, 50).

29 Como ellos mismos admiten en el prólogo de la 1ª edición, "la explicación de de la jerarquía de los tipos en la Introducción difiere un poco de la dada en el *12 en el cuerpo de la obra" (PM, p. vii). Sigo de preferencia esta última. Church (1976, 1984) ofrece una reconstrucción rigurosa de la teoría ramificada de los tipos. Para que el lector se haga una idea de lo que ello entraña, resumo en esta nota sus principales reglas de sintaxis. La clasificación en órdenes y tipos —que Church llama r-tipos, para distinguirlos de los simples tipos de la teoría simple— se aplica en primer lugar a las variables, en segundo lugar a las constantes con que es lícito sustituirlas. Los r-tipos están agrupados en lo que Church llama *niveles* (*levels*), cada uno de los cuales se conoce por un número natural. Dichos niveles son acumulativos: el recorrido de una variable de cierto r-tipo incluye los recorridos de todas las variables de nivel inferior al suyo. El nivel 1 corresponde a las variables cuyo recorrido sólo contiene lo que Whitehead y Russell llamaban funciones predicativas. La escritura conceptual adoptada comprende los símbolos lógicos de negación, disyunción y cuantificación universal, paréntesis, una lista infinita de variables de cada r-tipo y listas finitas —posiblemente vacías— de constantes de distinto r-tipo. Hay un tipo *i* integrado por las variables individuales. Si $\beta_1, \ldots, \beta_m$ son r-tipos ($m \geq 0$), hay un r-tipo $(\beta_1, \ldots, \beta_m)/n$ integrado por variables predicativas *m*-arias de nivel *n* ($n \geq 1$). El r-tipo $(\alpha_1, \ldots, \alpha_m)/k$ es inferior al r-tipo $(\beta_1, \ldots, \beta_m)/n$ si y sólo si $k < n$ y $\alpha_h = \beta_h$ $(1 \leq h \leq m)$. Si $\beta_1 = \ldots = \beta_m = i$, en vez de $(i, \ldots, i)$ escribimos *m*. El orden de una variable $\xi$ se define recursivamente así: (i) si $\xi$ es del r-tipo *i*, el orden de $\xi$ es 0; (ii) si $\xi$ es del r-tipo $(\beta_1, \ldots \beta_m)/n$ y *N* es el orden más alto que puede tener una variable de uno de los r-tipos $\beta_1, \ldots \beta_m$, el orden de $\xi$ es $N + n$. Una fórmula está bien formada (i) si consta sólo de una variable proposicional, esto es, una variable de alguno de los r-tipos $0/n$; (ii) si consta de una variable del r-tipo $(\beta_1, \ldots, \beta_m)/n$ ($m > 0$) seguida de *m* variables y la *h*-ésima variable

En aras de la fidelidad histórica, no me esforzaré en aclarar si las proposiciones, funciones proposicionales, variables libres y ligadas, y otros entes por el estilo, de que tendré que hablar, son expresiones lingüísticas u objetos suprasensibles. La necesidad de ser completamente inequívoco en este respecto, inculcada más tarde por Hilbert y sus discípulos, aparentemente no fue reconocida por Whitehead y Russell. Llamaré, con ellos, *matriz* a una función proposicional de $n$ variables ($n \geq 1$) que no contiene ninguna variable ligada. De una matriz $\phi\hat{x}_1 \ldots \hat{x}_n$ se derivan funciones que no son matrices, por generalización universal o existencial sobre una de sus variables; simbólicamente: $\forall x_k \phi\hat{x}_1 \ldots \hat{x}_{k-1} x_k \hat{x}_{k+1} \ldots \hat{x}_n$ y $\exists x_k \phi\hat{x}_1 \ldots \hat{x}_{k-1} x_k \hat{x}_{k+1} \ldots \hat{x}_n$ ($1 \leq k \leq n$).[30] Según Whitehead y Russell, "toda función posible que no sea una matriz se deriva de una matriz" por una o más generalizaciones (PM, I, 162).[31] Definiremos recursivamente el *orden* de una variable, matriz, fun-

---

($1 \leq h \leq m$) es del r-tipo $\beta_h$; (iii) si se obtiene reemplazando variables por constantes del mismo r-tipo en una fórmula bien formada como la descrita en (ii); (iv) si se construye a partir de fórmulas bien formadas por negación, disyunción o cuantificación. Además de las reglas de inferencia habituales de un cálculo predicativo de orden mayor que 1, Church adopta dos axiomas (esquemáticos) de comprensión:

> (I) $\exists p(p \leftrightarrow P)$, donde p es una variable del r-tipo $0/n$ que no figura como variable libre en P, todas las variables ligadas de P son de orden menor que $n$ y ninguna constante ni variable libre de P es de orden mayor que $n$.
> (II) $\exists f \forall x_1 \ldots \forall x_m (fx_1 \ldots x_m \leftrightarrow P)$, donde f es una variable predicativa del r-tipo $(\beta_1, \ldots \beta_m)/n$, $x_1, \ldots, x_m$ son variables diferentes de los r-tipos $\beta_1, \ldots \beta_m$, respectivamente, que pueden figurar como variables libres en P, todas las variables ligadas de P son de orden menor que el orden de f y ninguna constante ni variable libre de P es de orden mayor que el orden de f.

30 Whitehead y Russell hubiesen escrito
$(x_k).\phi(\hat{x}_1, \ldots, \hat{x}_{k-1}, x_k, \hat{x}_{k+1} \ldots, \hat{x}_n)$ en vez de $\forall x_k \phi\hat{x}_1 \ldots \hat{x}_{k-1} x_k \hat{x}_{k+1} \ldots \hat{x}_n$ y
$(\exists x_k).\phi(\hat{x}_1, \ldots, \hat{x}_{k-1}, x_k, \hat{x}_{k+1} \ldots, \hat{x}_n)$ en vez de $\exists x_k \phi\hat{x}_1 \ldots \hat{x}_{k-1} x_k \hat{x}_{k+1} \ldots \hat{x}_n$.

31 Olvidan, aparentemente, que una función puede derivarse también de *varias* matrices, combinando las dos formas de generalización con la negación y la disyunción. Tal olvido resulta inofensivo en el caso de las matrices de primer orden (que sólo contienen variables que designan ambiguamente *individuos*), pues, si $\phi\hat{x}_1 \ldots \hat{x}_n$ es una función proposicional de $n$ variables individuales formada a partir de varias matrices por cualquier combinación de generalizaciones, negaciones y disyunciones, hay una función proposicional "prenexa" $Qz_1 \ldots Qz_p \psi z_1 \ldots z_p \hat{x}_1 \ldots \hat{x}_n$, lógicamente equivalente a $\phi\hat{x}_1 \ldots \hat{x}_n$, y formada por generalizaciones sucesivas a partir de la matriz de primer orden $\psi\hat{z}_1 \ldots \hat{z}_p \hat{x}_1 \ldots \hat{x}_n$ (véase el Apéndice XIII; la Q representa indistintamente un signo de cuantificación existencial o universal).

ción o proposición. Diré que una variable es de orden 0 si los objetos ambiguamente designados por ella son individuos, y que es de orden $n$ ($n \geq 1$) si tales objetos son funciones de $n$-ésimo orden.[32] Una matriz se dice *de primer orden* si sólo contiene variables de orden 0. Una función se dice *de primer orden* si es una matriz de primer orden o se deriva de una matriz de primer orden por generalización. Si se ligan por generalización *todas* las variables libres de una matriz de primer orden, se obtiene una *proposición de primer orden*. Una matriz cuyas variables son todas de orden menor que $n > 1$ y que contiene por lo menos una variable de orden $n - 1$ se dice *de n-ésimo orden*. Una función se dice *de n-ésimo orden* si es una matriz de $n$-ésimo orden o se deriva de una matriz de $n$-ésimo orden por generalización. Si se ligan por generalización *todas* las variables libres de una matriz de $n$-ésimo orden, se obtiene una *proposición de n-ésimo orden*. Evidentemente, una función o proposición de $n$-ésimo orden sólo contendrá variables de orden menor que $n$ y por lo menos una variable de orden $n - 1$. Con exasperante exuberancia terminológica, Whitehead y Russell, en vez de *matriz*, suelen decir *función predicativa* (*vide supra*, p. 190, n. 20).† Para distinguir la representación simbólica de una función predicativa o matriz de la de otras funciones insertan un signo de admiración después de la minúscula griega inicial, por ejemplo, así: $\phi!x$, $\psi!(x,y)$. Este método de representación se extiende también a las proposiciones que se derivan de una matriz mediante la sustitución de todas sus variables por constantes (llamadas *proposiciones elementales* si la matriz en cuestión es de primer orden).

La teoría ramificada de los tipos resuelve todas las paradojas que hemos mencionado. A modo de ejemplo, cito dos de sus soluciones. Cuando el mentiroso dice "estoy mintiendo", o bien habla sin sentido —en cuyo caso no hay paradoja—, o bien dice que una proposición $p$ aseverada por él es falsa. Sea $p$ de orden $n$. Entonces, la proposición "estoy mintiendo", que se refiere a $p$, es de orden superior a $n$ y por lo tanto no puede ser idéntica a $p$ (Russell LK, p. 79).[33] Consideremos ahora la paradoja de Berry. Ella concierne a la

[32] Whitehead y Russell no asignan órdenes a la variables en la explicación que estoy parafraseando, pero en PM, I, 167 hablan, sin previo aviso, de una "variable ligada de orden $n$", en una acepción que no puede ser otra que la que defino en el texto.

[33] Whitehead y Russell (PM, I, 38) se valen de un razonamiento análogo para disolver un ridículo argumento contra el escepticismo que todavía suele escucharse. Conforme a ese argumento, quien asevera que no sabe *nada* se contradice, pues reclama saber *al*

**† LA CONFUSIÓN ES MÍA. RECTIFICO:** Una función predicativa es una función proposicional que no contiene variables ligadas de orden superior a sus variables libres. Una matriz, en cambio, no contiene variables ligadas. [R.T.-2005]

función proposicional '$x$ es el entero positivo más pequeño que no se puede describir en castellano con menos de dieciocho palabras'. Usaré la expresión simbólica '$\Psi\xi$' para decir que $\xi$ es una función proposicional de una variable expresable en castellano con menos de 18 palabras (además de la variable $x$ y la cópula 'es'). Escribiré N$x$ para decir que $x$ es un entero positivo y $x \leq y$ para decir que el entero $x$ es menor o igual que el entero $y$. Conforme al análisis de Russell (1905), la función involucrada en la paradoja de Berry puede entonces simbolizarse así:

$$\exists y(x = y \wedge \mathrm{N}y \wedge \forall\phi\hat{z}((\phi y \rightarrow \neg\, \Psi\phi\hat{z}) \wedge \forall w((\mathrm{N}w \wedge (\phi w \rightarrow \neg\, \Psi\phi\hat{z})) \rightarrow y \leq w)))$$

La variable ligada $\phi\hat{z}$ tiene necesariamente un orden determinado $n$, inferior, por cierto, al de la función en que figura. Por lo tanto, no es contradictorio afirmar en castellano que un cierto número $q$ satisface esta función, aunque ello entrañe describir a $q$ con menos palabras que las que se está diciendo que hacen falta para eso. Aunque $q$ no satisfaga ninguna función de orden $n$ expresable en castellano con menos de 18 palabras, bien puede satisfacer una función de orden mayor que $n$ expresable en castellano con sólo 17 palabras.[34]

La victoria de Russell y Whitehead sobre las paradojas tiene un precio que a muchos ha parecido excesivo. El inconveniente de la teoría de los tipos que más salta a la vista es que ella multiplica los entes más allá de todo lo que uno se hubiera nunca imaginado que podía ser necesario. Piénsese en la relación de identidad. Normalmente pensaríamos que la función proposicional $\hat{x} = \hat{y}$ genera una proposición verdadera si sus dos variables se reemplazan por nombres o descripciones de una misma cosa, cualquiera que sea, y genera una proposición falsa si se las reemplaza por nombres o des-

*menos* eso que asevera. Según la teoría de los tipos, la oración 'no sé nada' sólo puede expresar que no hay ninguna proposición de orden menor que cierto entero positivo $n$ que diga algo que yo sé. Pero el orden de la proposición así expresada es, entonces, por lo menos igual a $n$.

34 Las dos soluciones citadas y otras cinco más —correspondientes a las paradojas de Burali-Forti y Richard y dos formas de la paradoja de Russell— aparecen en Whitehead y Russell, PM, I, 62–64. Church 1976 da una formulación y solución rigurosas de la paradoja de Grelling citada en la nota 19.

cripciones de dos cosas distintas, cualesquiera que sean. Pero según la teoría de los tipos no puede haber una función proposicional cuyo ámbito de significación sea la totalidad de las cosas (ni, mucho menos, el producto cartesiano de dicha totalidad consigo misma). Tiene que haber por eso muchas relaciones de identidad diferentes, una para cada tipo. El símbolo '=', en que veíamos un paradigma de la univocidad, se nos revela así como infinitamente ambiguo. No menos ambiguo es el símbolo '$\Lambda$' con que Whitehead y Russell designan lo que llaman la *clase nula*, esto es, la clase $\hat{x}(x \neq x)$ que contiene a todo objeto $x$ que no es idéntico a sí mismo (PM, I, 216).[35] Análoga ambigüedad afecta a los números cardinales. *El cardinal de la clase K* —llámemosle $|K|$— es la clase de todas las clases similares a $K$, es decir, de todas las clases $X$ tales que hay una correspondencia biunívoca cuyo dominio es $X$ y cuyo codominio es $K$ (Russell LK, p. 96). Pero, como explico en la nota 35, una clase es sólo una manera de hablar de funciones proposicionales de cierto tipo y en el caso de $|K|$ dicho tipo depende del tipo de los objetos de las clases $X$ y $K$. Así, variando $X$, se obtienen infinitos significados de '$|K|$' y sólo excepcionalmente será correcto decir que $|K| = |X|$.[36] La matemática puede sin duda arreglárselas con este resultado, sobre todo

35 Cuando se habla de clases a propósito de *Principia Mathematica* hay que entender esto a la luz del *20 de ese libro, donde se expone una "Teoría General de las Clases" que "evita suponer que haya algo así como clases" (PM, I, 187). La expresión simbólica '$\hat{x}(\phi x)$' —léase: 'la clase de los $x$ tales que $\phi x$'— no tiene ningún significado por sí misma, sino sólo en contexto: en el *20.01, se introduce la expresión '$f(\hat{x}(\phi x))$' como abreviatura de

$$\exists\psi(f(\psi!\hat{z}) \wedge \forall x(\psi!x \leftrightarrow \phi x))$$

En otras palabras, la frase 'la clase de los $x$ tales que $\phi x$ tiene la propiedad $f$' constituye una manera cómoda de decir que existe una función predicativa $\psi$ que tiene la propiedad $f$ y es tal que, cualquiera que sea el objeto $x$ (del tipo pertinente), $x$ tiene $\psi$ si y sólo si $x$ tiene $\phi$. (El lector observará, de paso, que no fue sólo "en aras de la brevedad" que Whitehead y Russell adoptaron la notación inconsecuente que comenté en la nota 26; en la fórmula recién citada, el cuantificador existencial liga la variable simbolizada con $\psi!\hat{z}$ al lado izquierdo de la conjunción y con $\psi!$ al lado derecho).

36 Si se estipula que la función $\hat{x} = \hat{y}$, aunque "ambigua en cuanto al tipo", es en todos los casos una función de dos variables del mismo tipo, resulta que '$|K| = |X|$' por regla general ni siquiera tiene sentido, ya que, aunque los objetos en $K$ (por ejemplo, los dedos de mi mano derecha) estén en correspondencia biunívoca con los objetos en $X$, no tienen que ser del mismo tipo que éstos (que podrían ser, por ejemplo, las funciones proposicionales que se generan reemplazando la variable $y$ por el número de gra-

si, siguiendo el ejemplo de Whitehead y Russell, utiliza sin remilgos un solo símbolo para representar propiedades o relaciones análogas de distinto tipo. Hay, con todo, cierta ironía en que la prosecución de su programa logicista llevara a Russell a admitir que un símbolo como '7' designa simultáneamente infinitas clases diferentes. En efecto, Russell rechazó el estructuralismo de Dedekind —para quien, como vimos en la Sección 2.3.2, cualquier sistema simplemente infinito tiene derecho a llamarse '*la* serie numérica'— y patrocinó en cambio, como Frege, la definición de los números en términos puramente lógicos, porque "queremos números que sirvan para contar los objetos comunes, y esto requiere que nuestros números tengan un significado *determinado* (*definite*), y no meramente que posean ciertas propiedades formales" (Russell 1919, p. 10).

Mas la teoría de los tipos no sólo hace irrisoria una de las motivaciones que animaron el logicismo, sino que contradice directamente la pretensión de deducir todas las proposiciones de la matemática pura de un pequeño número de principios lógicos fundamentales. Se recordará que para demostrar la existencia del siguiente de cada número finito Frege *agregaba* cada número finito $n$ a la clase de $n$ miembros integrada por los predecesores de $n$, con lo cual formaba una clase cuyo número era justamente el siguiente de $n$ (véase la demostración de [F2], pp. 172-74). Pero en virtud de la teoría de los tipos, si $n = |K|$, donde $K$ es la clase integrada por ciertos objetos $a_1, \ldots, a_n$, $n$ es un objeto de otro tipo que $a_1, \ldots, a_n$ y no puede pertenecer junto con ellos al ámbito de significación de una misma función proposicional. Por lo tanto, los objetos $a_1, \ldots, a_n$ y $n$ no forman una clase $K'$ tal que $|K'| = (n + 1)$. Para salvar la aritmética bajo estas condiciones, hay que postular que, si $K$ es una clase finita de objetos de un tipo dado, existe siempre un objeto $a$ de ese mismo tipo que no está contenido en la clase $K$; entonces el número de la clase $K \cup \{a\}$ será precisamente el siguiente de $|K|$. Tal postulado equivale al axioma que Whitehead y Russell llaman *Infin ax* —por 'Infinity axiom'— y adoptan formalmente (PM, *120.03). Éste dice que si $\alpha$ es un "cardinal inductivo" (esto es, un número finito en el sentido de Frege), existe por lo menos una clase —del tipo en cuestión— que contiene precisa-

mos que pesa cada uno de los dedos de mi mano derecha en la expresión '$x$ es un número real mayor que $y$').

mente $\alpha$ elementos.[37] Whitehead y Russell no pretenden hacer pasar este axioma por un principio lógico. Según ellos, *Infin ax* —al igual que su versión del Axioma de Selección, *Mult ax* (*88.03)— es "una hipótesis aritmética que algunos juzgarán evidente por sí misma, pero que preferimos mantener como hipótesis, aduciéndola como tal [esto es, como antecedente de una implicación material—R.T.] cada vez que sea pertinente" (PM, II, 203).

Aparte de que destruye el sueño logicista que la inspiró en sus inicios, la teoría de los tipos amenaza la posibilidad misma de la matemática en su forma clásica. Como atinadamente señala Zermelo (1908, p. 524; citado en la p. 69), ésta abunda en conceptos definidos sin prestar ni la más mínima atención al alegado "principio del círculo vicioso". Un caso ejemplar, aludido por Zermelo y analizado por Weyl (1918, p. 23) y Ramsey (1931, p. 64), es el concepto del *supremo* o cota superior mínima de un conjunto de números reales, tan importante en la fundamentación clásica del análisis. La circularidad inherente a su definición puede exhibirse fácilmente partiendo de la caracterización abstracta del sistema de los reales como un cuerpo† arquimédico completo presentada en el Glosario. Decimos que un número real $\sigma_K$ es

[37] *120.03 designa con el nombre "Infin ax" a la aseveración ambigua en cuanto al tipo que dice que si $\alpha$ es un cardinal inductivo, existe un objeto $x$ tal que $x \in \alpha$. Para mayor claridad, supongamos que $\alpha$ es un cardinal del tipo más bajo posible (lo que digo a continuación puede luego transferirse *mutatis mutandis* a los tipos superiores). Esto significa que $x \in \alpha$ si y sólo si $x$ es una clase de individuos y $\alpha$ es la clase de todas las clases similares a $x$ (dos clases $a$ y $b$ son similares si son del mismo tipo y hay una correspondencia biunívoca entre $a$ y $b$; cf. PM, *73, especialmente *73.02 y *73.03). En particular, si llamamos, con Whitehead y Russell, $\iota`a$ a la clase $\hat{y}(y = a)$ cuyo único miembro es el objeto $a$ y $\Lambda$ a la clase nula o vacía, tenemos que el cardinal 0 es la clase de todas las clases similares a $\Lambda$ y el cardinal 1 es la clase de todas las clases similares a $\iota`x$ (donde $x$ es un cierto individuo, que evidentemente puede ser cualquiera sin afectar la identidad del objeto designado por '1'). Si $\alpha$ es un cardinal, $\alpha + 1$ es la clase de todas las clases similares a $z \cup \iota`x$, donde $z \in \alpha$ y $x \notin z$, y por ende también es un cardinal. Si $\beta = \alpha + 1$, digo que $\beta$ tiene con $\alpha$ la relación (+1) (simbólicamente: $\beta(+1)\alpha$). Digo que $\beta$ tiene con $\alpha$ la relación $(+1)^*$ (simbólicamente: $\beta(+1)^*\alpha$) si (i) $\beta(+1)\alpha$ ó (ii) hay un cardinal $\gamma$ tal que $\beta(+1)\gamma$ y $\gamma(+1)^*\alpha$. El cardinal $\alpha$ es *inductivo* si y sólo si $\alpha(+1)^*0$. Así pues la aseveración ambigua en cuanto al tipo que bautizamos *Inf ax*, referida a objetos del tipo más bajo posible, dice que si $\alpha$ es cualquier cardinal obtenido por adiciones sucesivas de un individuo a la clase vacía hay una clase no vacía de individuos cuyo cardinal es $\alpha$. Obviamente, esto equivale a decir que existe una provisión inagotable de individuos para ir generando clases de diferente numerosidad mediante adiciones sucesivas de un individuo a la clase vacía.

una cota superior del conjunto de reales $K$ si $\sigma_K$ es mayor o igual que todo $k \in K$. El supremo de $K$, denotado por sup $K$, es la menor de las cotas superiores de $K$. En otras palabras, si $\varnothing \neq K \subset \mathbb{R}$, sup $K$ es un número real tal que (i) $\forall x(x \in K \subset \mathbb{R} \rightarrow x \leq \sup K)$ y (ii) $\forall y(\forall x(x \in K \subset \mathbb{R} \rightarrow x \leq y) \rightarrow \sup K \leq y)$. La cláusula (ii) de la definición de sup $K$ envuelve, pues, una referencia a la totalidad de los objetos $y$ tales que $\forall x(x \in K \subset \mathbb{R} \rightarrow x \leq y)$, uno de los cuales se pretende que sea el propio sup $K$, conforme a la cláusula (i). Con todo, la doctrina de Whitehead y Russell se apreciará mejor si consideramos este ejemplo en el contexto de una de las interpretaciones clásicas de los números reales como conjuntos de racionales. Siguiendo a Ramsey, utilizaré una reformulación de la interpretación de Dedekind (1872) similar a la adoptada en *Principia Mathematica*. Sea $\langle \mathbb{Q},+,\times,0,1 \rangle$ el cuerpo de los racionales (de un tipo dado). Una *sección* de $\mathbb{Q}$ es un conjunto no vacío $r \subset \mathbb{Q}$ tal que, si $\xi \in r$ y $\zeta \in \mathbb{Q} \backslash r$, entonces $\xi < \zeta$. Si $u$ y $v$ son secciones de $\mathbb{Q}$, diremos que $u$ es mayor o igual que $v$ si hay un $\xi \in u$ mayor o igual que todo $\eta \in v$. Sea $\mathbb{R}$ el conjunto de todas las secciones de $\mathbb{Q}$. Se puede demostrar que la estructura $\langle \mathbb{R},\boldsymbol{+},\boldsymbol{\times},0,1 \rangle$ es un cuerpo arquimédico completo —y, por ende, una interpretación legítima de los reales— si se estipula que $0 = \{\xi \in \mathbb{Q}: \xi \leq 0\}$, que $1 = \{\xi \in \mathbb{Q}: \xi \leq 1\}$, y que, para todo $u, v \in \mathbb{R}$, $u \boldsymbol{+} v = \{\xi + \zeta: \xi \in u \wedge \zeta \in v\}$ y $u \boldsymbol{\times} v = \{\xi \times \zeta: \xi \in u \wedge \zeta \in v\}$. Sea $K \subset \mathbb{R}$ un conjunto no vacío tal que existe una cota superior de $K$, esto es, una sección de $\mathbb{Q}$ mayor o igual que cualquier elemento de $K$. En tal caso, la unión de $K$, $\bigcup K = \{\xi \in \mathbb{Q}: \exists x(\xi \in x \wedge x \in \mathbb{R})\}$, es una sección de $\mathbb{Q}$ que satisface la doble condición siguiente: (C1) para cada racional $\zeta$ contenido en un $z \in K$ hay un racional $\xi \in \bigcup K$ tal que $\xi$ es mayor o igual que $\zeta$, y (C2) si $x \in \mathbb{R}$ es una cota superior de $K$ hay un racional $\xi \in x$ mayor o igual que cualquier elemento de $\bigcup K$. Esto implica que $\bigcup K$ es una cota superior de $K$ menor que cualquier otra. Así pues hemos demostrado, al parecer, que si un conjunto no vacío $K \subset \mathbb{R}$ tiene una cota superior, $K$ tiene una cota superior mínima o *supremo* $\bigcup K \in \mathbb{R}$. Pero la caracterización de $\bigcup K$ viola manifiestamente el principio del círculo vicioso, por cuanto emplea la función proposicional $\exists x(\xi \in x \wedge x \in \mathbb{R})$ que hace referencia a la totalidad de $\mathbb{R}$. La teoría de los tipos (ramificada) salva esta dificultad negando que las secciones de $\mathbb{Q}$ formen una totalidad acabada. El conjunto $\mathbb{R}$ que hemos tomado como base de nuestro cuerpo de los reales puede comprender, sí, todas las secciones de $\mathbb{Q}$ caracterizables mediante fun-

ciones de cierto orden, pero la función empleada para caracterizar a $\bigcup K$ tiene que ser en todo caso de otro orden, de modo que $\bigcup K \notin \mathbb{R}$. Resulta, entonces, que $\bigcup K$ no forma sumas ni productos con los elementos de $\mathbb{R}$, ni entra en el ámbito de verdad de ninguna función proposicional satisfecha por esos elementos. Podremos siempre llamar a $\bigcup K$ 'el supremo de *K*' y reunirlo con los supremos —así entendidos— de los otros subconjuntos acotados de $\mathbb{R}$, en un cuerpo arquimédico completo diferente de $\langle\mathbb{R},+,\times,0,1\rangle$. Pero la existencia del supremo de cada conjunto acotado de reales no servirá al propósito con el cual se la demuestra en la fundamentación clásica del análisis.

Una corriente minoritaria pero tenaz de la matemática del siglo XX sostiene que en vista de esto hay que descartar la fundamentación clásica del análisis y procurarle otra diferente, constructivista (Weyl 1918, Lorenzen 1965, Bishop 1967) o intuicionista (Brouwer 1907, 1930; Weyl 1921), aunque ello entrañe renunciar a ciertos teoremas que nos habíamos habituado a reputar verdaderos. Pero Whitehead y Russell no militan en esa corriente. Aunque éste se preguntaba cuánto habría que sacrificar de la matemática clásica para resolver las paradojas (Russell 1906, en EA, p. 156, citado arriba), cuando publican *Principia Mathematica* creen poder rescatarla entera. Con ese propósito, adoptan el Axioma de Reducibilidad, en virtud del cual, si $\psi\hat{z}$ es una función proposicional de cualquier orden hay siempre una función predicativa $\phi!\hat{z}$ con el mismo ámbito de significación que $\psi\hat{z}$, tal que, para todo argumento idóneo $x$, $\phi!x$ si y sólo si $\psi x$.[38] Aplicado al ejemplo que examinamos en el párrafo anterior, esto significa que si una función proposicional de cualquier orden circunscribe una clase de racionales, existe una función del orden mínimo que tiene exactamente el mismo ámbito de verdad y por lo

38 Whitehead y Russell enuncian un Axioma de Reducibilidad para funciones de una variable y otro para funciones de dos variables (*12.1, *12.11), y señalan que no enunciarán otros porque esos dos son los únicos que utilizarán en su obra (PM, I, 167). Ambos son, por cierto, esquemas ambiguos en cuanto al tipo. En la escritura conceptual de PM, el esquema correspondiente a una función proposicional de *n* variables se enunciaría así:

$$\vdash:(\exists f):\phi(x_1,\ldots,x_n).\equiv_{x_1,\ldots,x_n}.f!(x_1,\ldots,x_n)$$

(Las variables escritas como subíndices a la derecha del símbolo de equivalencia '≡' deben leerse como cuantificadores universales, a la manera de Peano explicada en la p. 149 después de la fórmula 50.)

tanto circunscribe la misma clase. Así, aunque el supremo de un conjunto acotado de elementos del cuerpo arquimédico completo ℝ tenga que caracterizarse mediante una función proposicional que hace referencia a la totalidad de ℝ y, por lo mismo, no puede aplicarse a elementos de ℝ sin violar el principio del círculo vicioso, esto no importa: el Axioma de Reducibilidad nos asegura que si dicha función existe también hay otra que no hace referencia a la totalidad de ℝ y expresa una condición satisfecha por los mismos racionales que satisfacen la condición expresada por la primera, de modo que ella también define la misma sección de ℚ. El Axioma de Reducibilidad, claro está, no nos dice cuál es la función predicativa equivalente a una función no predicativa dada ni nos suministra un método para construirla. Más que un principio lógico o matemático, el Axioma de Reducibilidad parece la cruda expresión de un deseo. No es raro, entonces, que la seguridad que ofrece haya inspirado en general poquísima confianza, a pesar del ingenio invertido como ahora veremos en su defensa.

El siguiente argumento en pro del Axioma de Reducibilidad se publicó primero con la sola firma de Russell (1908, § V), pero reaparece en el *12 de *Principia Mathematica*. Supongamos por un momento que las clases son objetos reales y que cada función proposicional $\phi\hat{x}$ determina la clase $\{x{:}\phi x\}$ a la que pertenece cada objeto que está en el ámbito de verdad de $\phi\hat{x}$. El Axioma de Reducibilidad se deduce entonces fácilmente, ya que, bajo el supuesto indicado, si existe la función $\phi\hat{x}$ existe también la función $\psi!\hat{x}$ definida por la condición $(\psi!x \leftrightarrow x \in \{x{:}\phi x\})$, la cual evidentemente es predicativa. Pero, como sabemos, nuestro supuesto implica también la paradoja de Russell y tiene que ser rechazado. En cambio, el Axioma de Reducibilidad —hasta donde sabemos— no implica paradoja alguna.[39] Sería, según eso, estrictamente más débil que dicho supuesto y, por ende, es un candidato excelente para sustituirlo una vez que éste se probó demasiado fuerte. Como dicen nuestros autores:

> Para que sea posible la matemática, es absolutamente necesario [ . . . ] que dispongamos de un método para hacer aseveraciones que habitualmente serán equivalentes a lo que queremos decir cuando hablamos (inexacta-

39 Chwistek (1922) creyó haber derivado de él la paradoja de Richard, y Copi (1950) la de Grelling, pero estaban equivocados. Véase Copi 1971, pp. 96s.

mente) de "todas las propiedades de *x*". (Una "propiedad de *x*" puede definirse como una función proposicional satisfecha por *x*.) Por lo tanto, hay que hallar, en lo posible, un método para reducir el orden de una función proposicional sin afectar la verdad o falsedad de sus valores. Parece que esto es lo que el sentido común efectúa aceptando que hay *clases*. Dada una función proposicional $\psi x$ de cualquier orden, se supone que esto equivale, para todos los valores de *x*, a una aseveración de la forma "*x* pertenece a la clase $\alpha$". Ahora bien, suponiendo que exista un ente tal como la clase $\alpha$, esta aseveración es de primer orden, puesto que no alude a ninguna función variable. De hecho, su única ventaja práctica sobre la aseveración original $\psi x$ reside en que es de primer orden. No hay ninguna ventaja en suponer que realmente hay clases, y la contradicción sobre las clases que no son miembros de sí mismas muestra que, si hay clases, tienen que ser algo radicalmente diferente de los individuos. Parecería que el solo propósito al que sirven las clases y una razón capital que las hace lingüísticamente convenientes es que proveen un método para reducir el orden de una función proposicional. Por lo tanto, no supondremos nada de lo que parece entrañar la aceptación de las clases por el sentido común, excepto esto: que toda función proposicional es equivalente, para todos sus valores, a alguna función predicativa del mismo argumento o argumentos.

(Whitehead y Russell, PM, I, 166)[40]

40 Llama la atención la frase 'toda función proposicional' empleada en la última cláusula, ya que la generalización se refiere obviamente a una totalidad ilegítima. En el enunciado formal del Axioma de Reducibilidad en *12.1 no figura un cuantificador universal cuya variable ligada recorra todas las funciones proposicionales que comparten un dado ámbito de significación, pero cabe sostener que éste es un arreglo puramente cosmético. Como señala Wittgenstein, en una carta a Russell escrita probablemente en el verano de 1913:

> Tu axioma de reducibilidad es $\vdash :(\exists f):\phi x.\equiv_x.f!x$. ¿No es esto un disparate (*is this not all nonsense*)? Puesto que esta prop[osición] sólo tiene un significado si podemos convertir la $\phi$ en una variable *ligada* (*apparent*). Si no podemos hacerlo, nunca podrá deducirse ninguna ley general de tu axioma. El axioma mismo me parece en este momento sólo un truco de prestidigitador (*a mere juggling trick*). Díme por favor si encierra algo más. El axioma como lo habéis enunciado es sólo un esquema y la verdadera Pp debiera ser $\vdash :.(\phi):(\exists f):\phi x.\equiv_x.f!x$ y ¿para qué serviría eso (*what would be the use of that*)?
>
> (Wittgenstein 1974, p. 26)

Dicen en otro lugar que la razón para aceptar un axioma, “lo mismo que otra proposición cualquiera” sólo puede ser “inductiva”, a saber, “que muchas proposiciones que son casi indudables se deducen de él, que no se conoce otra vía igualmente plausible por la cual esas proposiciones podrían ser verdaderas aunque el axioma fuera falso, y que de él no se puede deducir nada que sea probablemente falso” (PM, I, p. 59). En el caso del Axioma de Reducibilidad, las consecuencias que alegadamente lo hacen verosímil son, ante todo, los teoremas del análisis. Cuando Whitehead y Russell publican estas palabras en 1910, el cuestionamiento intuicionista del análisis clásico —“la amenaza bolchevique de Brouwer y Weyl” contra las matemáticas (Ramsey 1931, p. 56)— había recién comenzado y estaba confinada aún a la tesis doctoral, en holandés, y otros trabajos juveniles de Brouwer (*vide* Apéndice XIX). Pero el mismo año 1908 en que Russell publicó el Axioma de Reducibilidad, Zermelo había trazado otra vía “por la cual esas proposiciones podrían ser verdaderas aunque el axioma fuera falso”, una que la gran mayoría de los matemáticos considera mucho más transitable que la teoría (ramificada) de los tipos combinada con el Axioma de Reducibilidad. Es interesante comparar las dos alternativas. Zermelo viene de la escuela de Cantor quien, como sabemos, nunca enseñó que cada predicado determinara un conjunto (ni que cada conjunto fuera caracterizable por un predicado). Desde este punto de vista, las paradojas no ocasionan una revolución del pensamiento, pero hacen muy aconsejable proceder con cuidado en la caracterización de conjuntos mediante condiciones necesarias y suficientes. Los axiomas de Zermelo (1908a), concebidos ante todo para hacer perspicua su demostración del Teorema del Buen Orden, regulan tales caracterizaciones en una forma que —una vez revisada por Skolem y Fraenkel (Secciones 1.8.2 y 1.8.3)— ha resultado generalmente adecuada a las necesidades de los matemáticos. La estrategia de Zermelo consiste en admitir como existentes sólo aquellos conjuntos que la experiencia matemática revela imprescindibles. Sigue, pues, lo que puede llamarse una vía *minimalista* y, por ende, genuinamente “inductiva”, en el sentido de Whitehead y Russell. *Maximalista*, en cambio, es la seguida por éstos al adoptar el Axioma de Reducibilidad, que postula la existencia de una función predicativa generalmente desconocida y probablemente inefable para *cada* función proposicional que concibamos (y también para las que no concebimos). Tales entes son presumiblemente más tenues que los conjuntos que postula Zermelo, pero existencia es

existencia y Whitehead y Russell tienen que acomodarlos de alguna manera en su ontología. No se sabe cómo. En cambio, me parece claro que, a pesar de la anunciada neutralidad con respecto a la existencia de clases, *Principia Mathematica* las admite sólo como una manera de hablar (del modo descrito en la nota 35). En efecto, si una clase —o un conjunto— existe de veras no puede hacerle mella que la caractericemos impredicativamente. En tal caso, pues, no puede justificarse la prohibición contenida en el principio del círculo vicioso.[41]

La Introducción a la segunda edición de *Principia Mathematica* (1927), escrita por Russell, propone varios cambios que el lector debe subentender en el cuerpo de la obra, reproducido de la primera edición con sólo pequeñas correcciones. Uno de los más importantes es la supresión del Axioma de Reducibilidad. Para hacerla viable, se adopta la tesis de Wittgenstein de que todas las funciones de proposiciones son funciones veritativas —analizables, por ende, exhaustivamente en términos de disyunciones y negaciones (o de la sola "función palote" introducida por Sheffer)[42]— y que una función sólo

[41] Criticando el primer enunciado del principio del círculo vicioso —'Lo que envuelve el *todo* de una colección no puede ser un miembro de esa colección'— dice Gödel que esto vale "sólo si las entidades en cuestión han sido construidas por nosotros":

> En este caso es claro que tiene que haber una definición (a saber, la descripción de la construcción) que no se refiere a una totalidad a la que pertenece el objeto definido, pues la construcción de una cosa ciertamente no puede basarse en una totalidad de cosas a la que pertenezca la cosa misma que ha de ser construida. Sin embargo, si se trata de objetos que existen independientemente de nuestras construcciones, entonces no hay nada absurdo en la existencia de totalidades que contengan miembros que sólo puedan ser descritos (esto es, caracterizados unívocamente) por referencia a esa totalidad.
>
> (Gödel 1944, p. 136s.)

Conviene leer toda la crítica al principio del círculo vicioso en Gödel 1944, pp. 133–37 (trad. castellana en Gödel, OC, pp. 322–26).

[42] La función palote (*stroke function*) de Sheffer, llamada así porque se la representa mediante el signo |, es la función veritativa binaria definida por las condiciones (i) $p|q$ es verdadera si $p$ y $q$ son ambas falsas, y (ii) de otro modo, $p|q$ es falsa. Como evidentemente $\neg p$ si y sólo si $p|p$ y $p \vee q$ si y sólo si $(p|q)|(p|q)$, es claro que cualquier función veritativa puede caracterizarse mediante combinaciones de palotes. Las proposiciones generales se reducen a funciones veritativas si entendemos, con Wittgenstein, que la generalización universal $\forall x \phi x$ equivale a la conjunción $\phi x_1 \wedge \phi x_2 \wedge \ldots$ (y la generalización existencial $\exists x \phi x$ a la disyunción $\phi x_1 \vee \phi x_2 \vee \ldots$), donde las $x_i$ recorren todo

puede figurar en una proposición a través de sus valores. Como consecuencia de esto, si dos funciones proposicionales son satisfechas por exactamente los mismos argumentos, ellas mismas satisfacen exactamente las mismas funciones.[43] La renuncia al Axioma de Reducibilidad invalida, por cierto, algunas de las demostraciones del libro. La nueva sección *89 sobre la inducción matemática (Apéndice B), está destinada a reconstruir las bases de la teoría de los números naturales sin Axioma de Reducibilidad. Gödel anota un error en la demostración del lema principal *89.16, y concluye por eso que "la cuestión de si la teoría de los enteros puede obtenerse sobre la base de la jerarquía ramificada debe considerarse como no resuelta hasta la fecha" (1944, p. 146). En todo caso, aunque se reparara ese error, seguiría pendiente la reconstrucción de la teoría de los números reales, que ni siquiera se aborda en la segunda edición de *Principia*. Gödel considera además que la práctica de admitir como argumentos de un predicado de funciones, a funciones de orden más alto que el predicado mismo —que Russell defiende en la nueva Introducción (PM, I, xl y xli) y ejerce profusamente en el Apéndice B— significa que "el principio del círculo vicioso para las funciones proposicionales ha sido virtualmente descartado" (Gödel 1944, p. 134).

En la Introducción citada Russell agradece vivamente la colaboración de Frank Ramsey (PM, I, xiii, nota), por lo cual pienso que es posible que la vigorosa oposición del joven filósofo al Axioma de Reducibilidad haya contribuido a la decisión de abandonarlo. Sorprende, empero, que Russell no aluda a la clasificación de las paradojas que Ramsey publicó en 1925 y que según él hace posible eliminar la ramificación de los tipos (y con ella el motivo por el cual dicho Axioma se introdujo). Según Ramsey hay una clara diferencia entre paradojas tales como las de Russell, Cantor y Burali-Forti (que Russell ya tiene presentes en 1903), y las paradojas de Richard, Berry y el mentiroso (que Russell unió a las otras desde su 1906a, probablemente influido por Poincaré). Las paradojas del primer grupo "envuelven únicamente términos lógicos o matemáticos como *clase* y *número*, e indican que tiene

---

el universo del discurso. Como señala Gödel (1944, p. 144), esta concepción sólo está exenta de dificultades si el número de los individuos y de los predicados primitivos es finito.

43 Russell resume esta conclusión así: "Según nuestra presente teoría todas las funciones de funciones son extensionales, es decir, $\phi x \equiv_x \psi x . \supset . f(\phi\hat{x}) \equiv f(\psi\hat{x})$" (PM, I, xxxix).

que haber una falla en nuestra lógica o en nuestra matemática"; pero las del segundo grupo —al cual también pertenece la paradoja de Grelling citada en la nota 19— "no pueden formularse empleando sólo términos lógicos, puesto que todas contienen alguna referencia al pensamiento, el lenguaje o el simbolismo, que no son términos formales, sino empíricos" (Ramsey 1931, p. 20). Para resolver las paradojas del primer grupo basta la jerarquía simple de los tipos. La jerarquía ramificada se necesita en *Principia Mathematica* sólo para evitar las del segundo grupo, pero Ramsey muestra que este método de solución no es imprescindible. La clasificación de Ramsey ha sido generalmente aceptada, y es corriente llamar *paradojas sintácticas* a las del primer grupo y *paradojas semánticas* a las del segundo. Explicaré esta nomenclatura en el Capítulo 3.1, cuando hable de la solución de las paradojas semánticas propuesta por Tarski.[44]

[44] El distingo entre los dos grupos de paradojas está implícito ya en la observación que hizo Peano (1906) a propósito de la paradoja de Richard: ella no pertenece a la matemática sino a la lingüística.

## 2.5 ARITMÉTICA FINITISTA

En los próximos capítulos examinaré varias contribuciones a la ejecución del programa de Hilbert, publicadas antes de que Kurt Gödel pusiera en evidencia la imposibilidad de llevarlo a cabo conforme a su concepción inicial (Gödel 1930a, 1931; cf. Capítulo 2.1, *ad finem*). Pero antes de abordarlas, me referiré al trabajo de Thoralf Skolem, "La fundamentación de la aritmética mediante el modo recursivo de pensar" (1923). Aunque fue concebido al margen de la escuela de Hilbert, será reconocido por ésta como un paradigma de la matemática finita, de los conceptos y procedimientos sustantivos a utilizarse en la vindicación del sistema formal de la matemática clásica del infinito.[1] En su Observación Final, Skolem explica que redactó el trabajo en 1919, después de estudiar *Principia Mathematica* (Skolem, SWL, p. 187). Durante ese estudio se le ocurrió que una buena parte de las matemáticas se podía fundamentar sin valerse de generalizaciones universales o existenciales con un campo de aplicación ilimitado. Como las paradojas de la lógica y la teoría de conjuntos obviamente suponen dicha forma de generalización, al evitársela se elude también el complicado aparato adoptado por Whitehead y Russell para prevenirlas. Skolem llama la atención sobre la idea de "aseveración funcional" (*Funktionalbehauptung*), que atribuye a Whitehead y Russell y caracteriza así: "Una aseveración funcional consistirá en afirmar una proposición (*Aussage*) como válida en un caso que se deja indeterminado" (Skolem, SWL, p. 153).[2] Con su trabajo se propone mostrar lo siguiente: "*Si*

1 Sobre el distingo entre matemática sustantiva y formal, véanse las citas de Hilbert y von Neumann en las pp. 123 y 124, notas 13 y 14.

2 A continuación traduzco el pasaje de *Principia Mathematica* de donde Skolem toma esta idea:

> Cuando aseveramos algo que contiene una variable libre (*a real variable*), en rigor no cabe decir que estemos aseverando una *proposición*, pues sólo obtenemos una proposición determinada al asignar un valor a la variable, y entonces nuestra aseveración sólo se aplica a un caso determinado y no tiene de ningún modo la misma fuerza que antes. Cuando lo que aseveramos contiene una variable libre

*concebimos las aseveraciones universales de la aritmética como aseveraciones funcionales y nos basamos en el modo recursivo de pensar, dicha ciencia se puede fundamentar consecuentemente sin aplicar los conceptos 'always' y 'sometimes' empleados por Russell y Whitehead*" (p. 153, cursiva de Skolem). El "modo recursivo de pensar" (*rekurrierende Denkweise*) consiste en valerse libremente de la inducción matemática para la definición de conceptos y la demostración de teoremas. Combinado con la susodicha idea de aseveración funcional, este modo de pensar no deja lugar a dudas. Sea $\varphi(x)$ la aseveración funcional que dice que un número indeterminado $x$ tiene la propiedad $\varphi$. Sea $\sigma n$ el siguiente del número $n$. Si podemos establecer la proposición $\varphi(1)$ y la aseveración funcional $\varphi(x) \to \varphi(\sigma x)$, podemos evidentemente concluir que $\varphi(x)$. En efecto, como quiera que se fije el número $x$, la conclusión se obtiene mediante una serie finita de aplicaciones de la regla *modus ponens*. Análogamente, para definir una función $f$ con argumentos y valores numéricos basta estipular cuál es el número $f(1)$ y cómo se calcula $f(\sigma x)$ dado $f(x)$. Entonces podremos determinar el valor de $f$ correspondiente a un número determinado cualquiera mediante una serie finita de aplicaciones de dicho procedimiento de cálculo. Así pues, como Skolem proclama con orgullo, su reconstrucción de la aritmética es "consecuentemente finitista" y se basa en "el principio de Kronecker según el cual una determinación matemática sólo es una determinación efectiva si lleva a la meta al cabo de un número finito de intentos (*Versuche*)" (SWL, p. 188).

La teoría creada por Skolem (1923) se conoce como aritmética primitivo-recursiva y tiene ciertas limitaciones. Como señala van Heijenoort (1967, p. 302), en ella se puede decir y probar que existen infinitos números primos,

---

estamos aseverando una proposición completamente indeterminada, de entre todas las que resultan al darle diversos valores a la variable. Es conveniente decir que tales aseveraciones *aseveran una función proposicional*. Las fórmulas matemáticas corrientes contienen tales aseveraciones; por ejemplo, '$\text{sen}^2 x + \cos^2 x = 1$' no asevera este o aquel caso particular de la fórmula, ni asevera que la fórmula vale para *todos* los valores posibles de $x$, aunque equivale a esta última aseveración; simplemente asevera que la fórmula vale, dejando $x$ completamente indeterminado; y puede hacerlo legítimamente porque, como quiera que se determine $x$, resultará una proposición verdadera.

(Whitehead y Russell, PM, I, p. 18)

esto es, que existe, para cada número primo dado *x*, un número primo *z* tal que $x < z \leq x! + 1$; pero no sabríamos expresar que existen infinitos pares de primos gemelos —esto es, de la forma $\langle x, x+2\rangle$— mediante una fórmula que no contenga variables ligadas de recorrido ilimitado.

El trabajo de Skolem es a la vez largo y conciso y no puedo resumirlo aquí. Pero dado el carácter paradigmático que tiene para Hilbert y los suyos, es importante que el lector se forme una idea clara de sus métodos, que ilustraré con varios ejemplos. Antes de entrar en materia, Skolem hace algunas advertencias que conviene repetir. Ante todo, nos dice que concibe todas las funciones como "propiamente descriptivas". Las funciones proposicionales se distinguen sólo porque no admiten más valores que *verdadero* y *falso*. Específicamente concibe tales funciones descriptivas "como nombres propios funcionales, es decir, nombres propios cuya denotación depende de la elección de una o más variables". Según esto, $\sigma n$ se concibe como el nombre propio de un número, pero de tal suerte que el número así designado varía según como se elija el número *n*.[3] Advierte además que el signo de igualdad (=) entre dos expresiones indica siempre que ambas designan lo mismo; por lo tanto, cuando figura entre dos funciones proposicionales expresa —como en Frege— la equivalencia de éstas. Por último, anuncia que dará por supuestos los conceptos de *número natural* y de *el siguiente de un número*, así como el modo recursivo de pensar.

El § 1 de Skolem 1923 trata de la adición. La presenta como una función binaria definida recursivamente así: [$\text{i}_+$] $x + 1 = \sigma x$; [$\text{ii}_+$] $x + \sigma y = \sigma(x + y)$. Como Skolem —a diferencia de Peano— no menciona el número 1 entre los conceptos que da por supuestos, debemos entender que la estipulación [$\text{i}_+$] sirve en parte para caracterizarlo como el único número cuya suma con un número cualquiera *x* es el siguiente de *x*. (La existencia de tal número único estaría entonces implícita en la definición de la adición y sería un requisito de su viabilidad). La estipulación [$\text{i}_+$] autoriza a Skolem a escribir '$x + 1$' por 'el siguiente de *x*' (en vez de $\sigma x$, una notación que Skolem no

[3] Skolem comenta brevemente la teoría de las descripciones de Russell, según la cual una expresión como 'el autor de *Waverley*' es un símbolo incompleto que nada significa por sí solo. Dice que no le parece incuestionable y que, en todo caso, aunque fuese correcta aplicada a las descripciones del lenguaje cotidiano, no hay por qué concebir de este modo a las funciones descriptivas de la aritmética (SWL, p. 154).

utiliza). Así, la estipulación [$ii_+$] aparece escrita de este modo: $x + (y + 1) = (x + y) + 1$. Como los conceptos de *número natural* y *siguiente de un número* se han dado por supuestos, la estipulación [$i_+$] determina el valor de la suma de *cualquier* número $x$ y el número fijo 1. Por su parte, la estipulación [$ii_+$] nos enseña a calcular en un número finito de pasos la suma de $x$ y otro número cualquiera $y$, a partir del valor de $x + 1$ establecido por [$i_+$]. Skolem demuestra luego que la adición es asociativa, esto es, que, cualesquiera que sean los números $a$, $b$ y $c$, $a + (b + c) = (a + b) + c$. La argumentación se presenta con todo detalle como ejemplo típico de una prueba por inducción. La estipulación [$ii_+$] relativa a la adición significa que esta ecuación es válida si $c = 1$, cualesquiera que sean los números $a$ y $b$. Supongamos, entonces, que, cualesquiera que sean $a$ y $b$, ella vale para un número fijo $c$. De este supuesto se deduce fácilmente que la ecuación vale también si reemplazamos $c$ por el siguiente de $c$. En efecto, $a + (b + (c + 1)) = a + ((b + c) + 1)$, según la estipulación [$ii_+$]. Pero $a + ((b + c) + 1) = (a + (b + c)) + 1$, por la misma estipulación [$ii_+$]. Con arreglo a nuestro supuesto, $(a + (b + c)) + 1 = ((a + b) + c) + 1$. Pero $((a + b) + c) + 1 = (a + b) + (c + 1)$, según la estipulación [$ii_+$]. Hilvanando estas ecuaciones, tenemos pues que $a + (b + (c + 1)) = (a + b) + (c + 1)$, Q.E.D. El lector que haya visto el Apéndice XI reconocerá aquí la demostración de Dedekind. Lo mismo cabe decir del teorema de conmutatividad de la adición que Skolem demuestra enseguida.

Luego Skolem define la relación '$a < b$' ('$a$ es menor que $b$'), también mediante dos estipulaciones: [$i_<$] $\neg(a < 1)$;[4] y [$ii_<$] $(a < b + 1) \leftrightarrow ((a < b) \vee (a = b))$. Combinadas con la definición de suma, estas estipulaciones implican que 1 no es el siguiente de ningún número. En efecto, supongamos que $n$ es un número tal que 1 es el siguiente de $n$. Entonces, en virtud de [$ii_<$], $n < n + 1 = 1$; pero esto contradice la estipulación [$i_<$]. Es fácil demostrar la tricotomía: cualesquiera que sean los números $a$ y $b$, $(a < b) \vee (a = b) \vee (b < a)$. En vez de $(a < b) \vee (a = b)$, escribiremos $(a \leq b)$. Omito la definición de la multiplicación y la prueba de que es asociativa, conmutativa y distributiva porque —como en el caso de la adición— son prácticamente iguales a las de Dedekind (Apéndice XI).

Para definir la relación '$a$ es divisible por $b$' —simbolizada $D(a,b)$—

4 Uso el simbolismo de nuestro cálculo predicativo (Apéndice IX), en vez de los símbolos lógicos tomados de Schröder que Skolem emplea.

Skolem introduce el importante concepto de cuantificación sobre un dominio acotado. Si $\varphi(x)$ es una aseveración funcional cualquiera y $n$ es un número determinado, en vez de la conjunción $\varphi(1) \wedge \varphi(2) \wedge \ldots \wedge \varphi(n)$ escribimos $(\forall x{\leq}n)\varphi(x)$ —léase: 'para todo $x$ menor o igual que $n$, $\varphi(x)$'— y en vez de la disyunción $\varphi(1) \vee \varphi(2) \vee \ldots \vee \varphi(n)$ escribimos $(\exists x{\leq}n)\varphi(x)$ —léase: 'hay un $x$ menor o igual que $n$, tal que $\varphi(x)$'. Los cuantificadores '$(\forall x{\leq}n)$' y '$(\exists x{\leq}n)$' ligan variables de recorrido finito y las fórmulas que los contienen pueden siempre reemplazarse por la correspondiente conjunción o disyunción. Su uso no infringe, pues, la decisión inicial de no emplear 'siempre' (*always*) o 'a veces' (*sometimes*) a la manera de Whitehead y Russell. La divisibilidad puede ahora definirse sencillamente así:

$$D(a,b) \leftrightarrow (\exists x{\leq}a)(a = bx)$$

A continuación se abordan la subtracción y la división. A propósito de ellas, Skolem introduce la idea de función parcial: si $f$ es una función de este tipo, $fx$ designa un cierto número sólo si $x$ es uno de los números pertenecientes a un dominio acotado expresamente. Atendamos a sus explicaciones. La subtracción se define habitualmente mediante la estipulación $(z - y = x) \leftrightarrow (z = x + y)$. Según esto, obviamente, la *diferencia* $z - y$ no existe si $z \leq y$, puesto que en tal caso es imposible que $z = x + y$. Pero puede demostrarse que hay un número igual a $z - y$ cuando quiera que $y < z$. Tradicionalmente, este aserto se habria enunciado así:

$$\neg(y < z) \vee \exists x(x + y = z)$$

Pero esta fórmula con variables ligadas de recorrido ilimitado se puede ahora reemplazar con esta otra

$$\neg(y < z) \vee (\exists x{\leq}z)(x + y = z)$$

que Skolem demuestra. La división es objeto de un tratamiento análogo.

El trabajo culmina en los §§ 6 y 7, donde Skolem, sin recurrir a otros medios que los que ha declarado admisibles, define los conceptos de máximo común divisor, mínimo común múltiplo y número primo, y demuestra que si $1 < x$, el número $x$ se puede siempre representar de una y sólo una

manera como producto de números primos.[5] En estos dos parágrafos, Skolem a veces define un predicado $U(x)$, para cualquier número $x$, estipulando las condiciones necesarias y suficientes para que [i] $U(1)$ y [ii] $U(x)$ dado que $(\forall z<x)U(z)$. Este procedimiento es característico de la llamada definición por inducción transfinita (Apéndice VIII), pero Skolem explica que si $x$ es un número natural, tenemos que habérnoslas sólo con una variante de la definición recursiva ordinaria. Tenemos, en efecto, que el valor de la función proposicional $U(z) \wedge (z \leq x)$ para cualquier par de números $\langle x,z \rangle$ se conoce si se conoce el valor de $U(z)$ para $z$, y que, por otra parte, se sabe si $U(z)$ es verdadero o falso para un cierto $z \leq x$ si se conoce el valor de $U(z) \wedge (z \leq x)$ para ese $z$ y ese $x$. Por lo tanto, para definir $U(x)$ bastará definir $U(z) \wedge (z \leq x)$. El valor de $U(z) \wedge (z \leq x)$ puede definirse por inducción ordinaria. Estipulamos primero el valor de $U(1)$. Esto nos da el valor de $U(z) \wedge (z \leq 1)$ para cualquier $z$, ya que, si $z \neq 1$, $U(z) \wedge (z \leq 1)$ es falso. Luego suponemos establecido el valor de $U(z) \wedge (z \leq n)$ para cualquier número $z$, y estipulamos sobre esa base el valor de $U(z) \wedge (z \leq n + 1)$ para cualquier $z$. Ahora bien, si $n + 1 < z$, entonces $\neg(z \leq n + 1)$, de modo que $U(z) \wedge (z \leq n + 1)$ es falso; si $z < n + 1$, entonces $z \leq n$ y el valor de $U(z) \wedge (z \leq n)$ está ya determinado por hipótesis; por lo tanto, sólo hace falta estipular el valor de $U(z)$ para el caso en que $z = n + 1$. En otras palabras, "determinar del valor de $U(x+1)$ si esta función se supone conocida para cualquier $z < x + 1$ equivale a determinar el valor de la función proposicional $U(x) \wedge (z \leq x)$ para $x = n + 1$ y un número cualquiera $z$ si ya se lo conoce para $x = n$ y un número cualquiera $z$" (SWL, p. 172; he modificado algo la notación de Skolem).

Skolem también recurre a veces al modo de razonar propio de la inducción transfinita: De las premisas [α] $U(1)$ y [β] $(\forall z<x)U(z) \rightarrow U(x)$, inferir [γ] $U(x)$. Su utilización en el presente contexto está perfectamente justifica-

5 Skolem define la función binaria $a \blacktriangle b$ mediante la triple condición $(a = b \rightarrow a \blacktriangle b = a) \wedge (b < a \rightarrow a \blacktriangle b = (a - b) \blacktriangle b) \wedge (a < b \rightarrow a \blacktriangle b = a \blacktriangle (b - a))$ y muestra luego que $a \blacktriangle b$ es precisamente el máximo común divisor de $a$ y $b$. Asímismo, define la función $a \blacktriangledown b$ mediante la condición $a \blacktriangledown b = ab/(a \blacktriangle b)$ y muestra que $a \blacktriangledown b$ es el mínimo común múltiplo de $a$ y $b$. Para definir la función proposicional $P(x)$ —léase '$x$ es un número primo'— Skolem define una función proposicional binaria que designaré $\Pi(x,y)$ (Skolem utiliza también para ésta la misma letra $P$). Ésta queda caracterizada por la doble estipulación siguiente: [i] $\Pi(x,1)$; [ii] $\Pi(x,y + 1) \leftrightarrow (\Pi(x,y) \wedge (x = y + 1 \vee \neg D(x,y + 1)))$. Entonces, $P(x) \leftrightarrow (\Pi(x,y) \wedge (x \neq 1))$

da, por cuanto la premisa [β] implica siempre una aseveración que, combinada con [α], lleva a la conclusión [γ] por inducción matemática ordinaria. En efecto, si $x$ es un cierto número, la aseveración de que $(\forall z{<}x)U(z)$ equivale a la aseveración de que para ese $x$ y cualquier $z$, $(z < x) \to U(z)$. Por lo tanto, la premisa [β] $(\forall z{<}x)U(z) \to U(x)$ equivale a $((z < x) \to U(z)) \to U(x)$. Como $U(x) \to ((z = x) \to U(z))$, es claro que [β] implica que $((z < x) \to U(z)) \to ((z = x) \to U(z))$, lo cual equivale a $((z < x) \to U(z)) \to (((z < x) \to U(z)) \wedge ((z = x) \to U(z)))$. Pero $((z < x) \to U(z)) \wedge ((z = x) \to U(z))$ equivale a $((z < x) \vee (z = x)) \to U(z)$, y por consiguiente a $((z < x + 1) \to U(z))$. Por lo tanto, [β] implica que $((z < x) \to U(z)) \to ((z < x + 1) \to U(z))$. Ahora bien, de $U(1)$ y $((z < x) \to U(z)) \to ((z < x + 1) \to U(z))$ se infiere por inducción ordinaria que $(z < x) \to U(z)$ para cualquier número $x$. Reemplazando en este último aserto $x$ por $x + 1$ y $z$ por $x$, obtenemos la conclusión $U(x)$ por *modus ponens*.

Me referiré por último a un teorema intuitivamente obvio que podemos enunciar así: Si un cierto número tiene la propiedad $U$, hay un número que tiene la propiedad $U$ y que es menor que cualquier otro que la tenga. Las generalizaciones —una existencial y universal la otra— incluidas en la apódosis de este enunciado se pueden expresar sin pérdida de información mediante cuantificadores de dominio acotado. Para abreviar designo con $\mu xU(x)$ al número más pequeño que tiene la propiedad $U$ (Skolem emplea para ello la expresión ‘Min($U$,$x$)’). Es claro que si hay un cierto número $n$ tal que $U(n)$, entonces $\mu xU(x)$ —si existe— será igual o menor que $n$. Además, para informar que $\mu xU(x)$ es *menor* que cualquier *otro* número $z$ tal que $U(z)$ bastará decir que $\mu xU(x)$ es *igual* a cualquier $z \leq \mu xU(x)$ tal que $U(z)$. El contenido del teorema se deja, pues, formular mediante las dos aseveraciones siguientes:

[i] $$U(n) \to (\exists x{\leq}n)(\forall z{\leq}x)(U(x) \wedge (U(z) \to x = z))$$

[ii] $$((U(a) \wedge (\forall x{\leq}a)(U(x) \to x = a)) \wedge (U(b) \wedge (\forall y{\leq}b)(U(y) \to y = b))) \to a = b$$

[i] dice que $U(n)$ implica la existencia de un número que obedece a la descripción de $\mu xU(x)$ y [ii] dice que tal número es único. La aseveración [i] es obvia si $n = 1$. Supongamos, pues, que el teorema es verdad para cualquier

$n \leq m$ y que $U(m+1)$. Caben dos posibilidades: o bien hay un número $w < m + 1$ tal que $U(w)$, y entonces, por hipótesis, $(\exists x \leq w < m+1)(\forall z \leq x)(U(x) \wedge (U(z) \rightarrow x = z))$, de modo que $U(m+1) \rightarrow (\exists x \leq m+1)(\forall z \leq x)(U(x) \wedge (U(z) \rightarrow x = z))$; o bien no hay ningún número $w < m + 1$ tal que $U(w)$, en cuyo caso $m+1 = \mu x U(x)$ y es claro que $U(m+1) \rightarrow (\exists x \leq m+1)(\forall z \leq x)(U(x) \wedge (U(z) \rightarrow x = z))$. La aseveración [ii] se deduce fácilmente de la tricotomía de los números. Si $a \neq b$, entonces $a < b$ o $b < a$; pero si $b < a$ es claro que $(\forall x \leq a)(U(x) \rightarrow x = a)$ implica $\neg U(b)$ y si $a < b$, $(\forall y \leq b)(U(y) \rightarrow y = b)$ implica $\neg U(a)$. Así pues, $a \neq b$ implica que la prótasis de la aseveración [ii] es falsa, de modo que dicha aseveración es verdadera.

## 2.6 PRUEBAS DE CONSISTENCIA

### *2.6.1 Ackermann (1925)*

La tesis doctoral de Wilhelm Ackermann, "Fundamentación del 'tertium non datur' mediante la teoría de la no-contradicción de Hilbert", da testimonio de la confianza ciega que Hilbert y sus discípulos tenían en la viabilidad de su programa. Se trata de "borrar la fea mancha de las paradojas" del "edificio por otra parte tan perfecto de la matemática" y de "salvar partes valiosas de la matemática —por ejemplo, la teoría de conjuntos— del escepticismo que todo lo aniquila" (Ackermann 1925, p. 1). Con ese fin se propondrá un sistema axiomático suficiente para deducir la matemática entera y se buscará demostrar su consistencia (*Widerspruchsfreiheit*). Ahora bien, "las paradojas surgen en la matemática sólo allí donde se trata de totalidades infinitas, donde se emplean las palabras 'todos' y 'existe' y las formas de inferencia transfinita —como la 'inducción completa' y el 'tertium non datur'— que sólo pueden formularse mediante esas palabras" (p. 1). Obviamente, tales formas de inferencia deberán evitarse, so pena de circularidad, en las pruebas de consistencia.

> Hilbert evita estas dificultades con su conocido distingo entre matemática y metamatemática. La matemática se convierte en un acopio de fórmulas, de signos. También lo transfinito tiene aquí su figura visible, finita. La metamatemática suministra la prueba de consistencia para la matemática. Como la metamatemática sólo hace aseveraciones sobre cosas concretas, presentes a la vista, se las arregla enteramente sin inferencias transfinitas. Emplea sólo aquellas formas de inferencia primitivas y finitas que admiten aun los escépticos más tenaces.
>
> (Ackermann 1925, p. 1)

Ackermann formula la matemática axiomatizada en un cálculo predicativo cuyos símbolos lógicos, con una sola excepción que explicaré luego, corresponden exactamente a los de nuestro Apéndice IX. Para facilitar la lec-

tura los reemplazaré, pues, por éstos. Hay variables individuales *a*, *b*, *c*, . . . ; variables funcionales *n*-arias (para cada entero positivo *n*) *f*, *m*, . . . ; variables proposicionales *A*, *B*, *C*; variables predicativas *A*(*a*), *B*(*a*,*b*), . . . , $A_a f(a)$ (donde el subíndice *a* indica que *A* depende de la variable funcional *f* y no de su argumento *a*), . . . ; constantes individuales 0 y 1, y constantes funcionales. Estas últimas incluyen los símbolos + y × de la adición y la multiplicación, que se escriben entre los signos de los argumentos, y letras minúsculas griegas φ, ψ, . . . que simbolizan funciones *n*-arias (para diversos valores del entero positivo *n*) y se escriben delante del *n*-tuplo de los argumentos. Como variables metalingüísticas —esto es, como parte del alemán en que él escribe (o del castellano en que lo explico)— Ackermann usa letras góticas: las mayúsculas $\mathfrak{A}$, $\mathfrak{B}$, $\mathfrak{C}$, . . . representan fórmulas; las minúsculas $\mathfrak{a}$, $\mathfrak{b}$, $\mathfrak{c}$, . . . expresiones de cierto género que, siguiendo a Hilbert, llama *funcionales* (*Funktionale*). Éstas se pueden caracterizar así: toda combinación sintácticamente admisible de signos 0, 1, +, × y variables individuales es un funcional; si φ es una constante funcional *n*-aria y $\mathfrak{a}_1, \dots, \mathfrak{a}_n$ son funcionales, entonces $\varphi(\mathfrak{a}_1, \dots, \mathfrak{a}_n)$ es un funcional. Conforme a la tradición iniciada por Frege, se adoptan sólo dos reglas de inferencia. La primera es *modus ponens*: de las premisas $\mathfrak{S}$ y $\mathfrak{S} \rightarrow \mathfrak{T}$ inferir la conclusión $\mathfrak{T}$. La segunda es esta *regla de sustitución*: si $\mathfrak{A}$ es un axioma y *x* es una variable cualquiera, inferir de $\mathfrak{A}$ la fórmula que se obtiene reemplazando uniformemente *x* en $\mathfrak{A}$ por un funcional determinado si *x* es una variable individual, por una cierta función *n*-aria si *x* es una variable funcional *n*-aria, por una fórmula cualquiera si *x* es una variable proposicional del tipo *A*, *B*, . . . , por una fórmula que contenga una función 1-aria si *x* es una variable proposicional del tipo $A_a$ *f*(*a*), etc.

Los funcionales de la forma 0, 0 + 1, . . . , 0 + 1 + 1 + 1 + 1, . . . se llaman *numerales* (*Zahlzeichen*). Una expresión que sólo incluya numerales y uno o más de los símbolos lógicos =, →, ∧, ∨, ¬, es una *fórmula numérica*. Una fórmula numérica de la forma $\mathfrak{a} = \mathfrak{b}$ se dice *correcta* (*richtig*) si $\mathfrak{a}$ y $\mathfrak{b}$ son numerales iguales; de otro modo se dice *incorrecta* (*falsch*). Estipulamos que, si $\mathfrak{R}$ designa una fórmula numérica correcta y $\mathfrak{F}$ una incorrecta, las fórmulas $\mathfrak{R} \rightarrow \mathfrak{R}$, $\mathfrak{F} \rightarrow \mathfrak{R}$, $\mathfrak{F} \rightarrow \mathfrak{F}$, $\mathfrak{R} \wedge \mathfrak{R}$, $\mathfrak{R} \vee \mathfrak{R}$, $\mathfrak{R} \vee \mathfrak{F}$, $\mathfrak{F} \vee \mathfrak{R}$ y $\neg\mathfrak{F}$ son correctas y las fórmulas $\mathfrak{R} \rightarrow \mathfrak{F}$, $\mathfrak{R} \wedge \mathfrak{F}$, $\mathfrak{F} \wedge \mathfrak{R}$, $\mathfrak{F} \wedge \mathfrak{F}$, $\mathfrak{F} \vee \mathfrak{F}$ y $\neg\mathfrak{R}$ son incorrectas. Es claro pues que toda fórmula numérica es correcta o incorrecta. No hay otra alternativa. Como subraya Ackermann (1925, p. 4),

el *tertium non datur* tiene aplicación aquí en un sentido finito, intuitivo.

Ackermann enuncia el siguiente criterio de consistencia: un sistema axiomático que emplee el formalismo descrito se reputa consistente si y sólo es imposible deducir de sus axiomas la fórmula numérica $\neg(0 = 0)$. El método propuesto para demostrar la consistencia de un sistema dado $\Sigma$ puede describirse sumariamente como sigue: Sea $\langle \mathfrak{A}_1, \ldots, \mathfrak{A}_n \rangle$ una lista de fórmulas tales que $\mathfrak{A}_n$ es una fórmula numérica y para cada entero positivo $k \leq n$, $\mathfrak{A}_k$ es un axioma de $\Sigma$, o se infiere por sustitución de una fórmula que la precede en la lista, o se infiere por *modus ponens* de dos fórmulas que la preceden en la lista. Se da una regla para reemplazar por una fórmula numérica a cada fórmula de la lista que no sea una fórmula numérica. Se muestra que aplicando esa regla cualquier lista que cumpla con los requisitos estipulados se convierte en una lista de fórmulas numéricas correctas. Como la fórmula $\neg(0 = 0)$ es incorrecta y toda deducción de una fórmula numérica $\mathfrak{A}_n$ a partir de los axiomas de $\Sigma$ consiste en una lista $\langle \mathfrak{A}_1, \ldots, \mathfrak{A}_n \rangle$ de fórmulas que reúne los requisitos indicados, el procedimiento descrito hace evidente que la fórmula $\neg(0 = 0)$ no puede deducirse de los axiomas de $\Sigma$. Por lo tanto, $\Sigma$ es consistente.

Los axiomas para la matemática clásica propuestos por Ackermann forman dos grupos: axiomas para la matemática finita y "axiomas transfinitos". Ackermann da una demostración breve pero completa de la consistencia del primer grupo y luego bosqueja larga y tortuosamente una demostración análoga aplicable a todo el sistema. No puedo dar aquí más que algunas indicaciones incompletas sobre esta última. En cambio, presentaré la primera en detalle, para que el lector se forme una idea precisa de las intenciones y recursos de la metamatemática hilbertiana. (También Ackermann ofrece esa demostración sobre todo a título ilustrativo).

Los axiomas de la matemática finita son 16. Los Axiomas 1–12 bastan para deducir por sustitución y *modus ponens* todas las tautologías; los Axiomas 13 y 14 gobiernan el signo de identidad, y los Axiomas 15 y 16 son propiamente aritméticos. Los doy aquí en nuestra notación:

1. $A \rightarrow (B \rightarrow A)$
2. $(A \rightarrow (A \rightarrow B)) \rightarrow (A \rightarrow B)$
3. $(A \rightarrow (B \rightarrow C)) \rightarrow (B \rightarrow (A \rightarrow C))$
4. $(B \rightarrow C) \rightarrow ((A \rightarrow B) \rightarrow (A \rightarrow C))$
5. $(A \wedge B) \rightarrow A$
6. $(A \wedge B) \rightarrow B$
7. $A \rightarrow (B \rightarrow (A \wedge B))$
8. $A \rightarrow (A \vee B)$
9. $B \rightarrow (A \vee B)$
10. $(A \rightarrow C) \rightarrow ((B \rightarrow C) \rightarrow ((A \vee B) \rightarrow C))$
11. $A \rightarrow (\neg A \rightarrow B)$
12.[1] $(A \rightarrow B) \rightarrow ((\neg A \rightarrow B) \rightarrow B)$
13. $a = a$
14. $a = b \rightarrow (A(a) \rightarrow A(b))$
15. $\neg(a + 1 = 0)$
16. $\neg(a = 0) \rightarrow (a = \delta(a) + 1)$

La constante funcional δ utilizada en el Axioma 16 se define recursivamente por la doble condición: $\delta(0) = 0$ y $\delta(a + 1) = a$. Las constantes + y × se definen también recursivamente así: $a + 0 = a$ y $a + (b + 1) = (a + b) + 1$; $a \times 0 = 0$ y $a \times (b + 1) = (a \times b) + a$. Ackermann admite que la matemática finita se enriquezca indefinidamente con nuevas funciones, siempre que éstas se definan según un esquema de recursión del cual las tres definiciones recién citadas son ejemplos sencillos. Dicho esquema se deja formular con

[1] El *tertium non datur,* $(A \vee \neg A)$, se deduce de los Axiomas 8, 9 y 12 por sustitución y *modus ponens*. En efecto, sustituyendo $B$ por $\neg A$ en el Axioma 8, obtenemos [α] $A \rightarrow (A \vee \neg A)$. Sustituyendo $B$ por $\neg A$ en el Axioma 9, obtenemos [β] $\neg A \rightarrow (A \vee \neg A)$. Sustituyendo $B$ por $(A \vee \neg A)$ en el Axioma 12, obtenemos [γ] $(A \rightarrow (A \vee \neg A)) \rightarrow ((\neg A \rightarrow (A \vee \neg A)) \rightarrow (A \vee \neg A)))$. Mediante dos aplicaciones de *modus ponens* a las premisas α, β y γ, inferimos $(A \vee \neg A)$.

toda generalidad así: Considérese una constante funcional $n$-aria $\varphi$ y funcionales $\mathfrak{a}(x_1, \ldots, x_{n-1})$ y $\mathfrak{b}(x_1, \ldots, x_{n+1})$ que contienen sólo $n - 1$ y $n + 1$ variables distintas, respectivamente ($n \geq 1$); entonces $\varphi$ se puede caracterizar mediante la doble condición siguiente:[2]

$$\begin{aligned} \varphi(0,b_1, \ldots ,b_{n-1}) &= \mathfrak{a}(b_1, \ldots ,b_{n-1}) \\ \varphi(a + 1,b_1, \ldots ,b_{n-1}) &= \mathfrak{b}(a,b_1, \ldots ,b_{n-1},\varphi(a,b_1, \ldots ,b_{n-1})) \end{aligned} \tag{R}$$

Demostraré, siguiendo a Ackermann, que la fórmula $\neg(0 = 0)$ no puede deducirse por sustitución y *modus ponens* de los 16 axiomas de la matemática finita, suplementados con definiciones recursivas de constantes funcionales conforme al esquema descrito. Sea $\mathcal{L} = \mathfrak{A}_1, \ldots, \mathfrak{A}_n$ una lista de fórmulas que constituye una deducción de la fórmula numérica $\mathfrak{A}_n$ a partir de dicho sistema de axiomas. Esto supone que, para cada entero positivo $k \leq n$, $\mathfrak{A}_k$ es uno de los axiomas 1–16, o se infiere de una o dos fórmulas precedentes de $\mathcal{L}$ por una de las reglas de inferencia. Supondremos, además, que toda fórmula que precede a $\mathfrak{A}_n$ en $\mathcal{L}$ se invoca una y sólo una vez para inferir otra fórmula de $\mathcal{L}$; por lo tanto, si uno de los axiomas se utiliza para deducir $r$ fórmulas de $\mathcal{L}$, $\mathcal{L}$ contiene $r$ copias de ese axioma, y si una fórmula que no es un axioma se aduce en la deducción de otras $r$ fórmulas de $\mathcal{L}$, $\mathcal{L}$ contiene $r$ copias de dicha fórmula y de toda la secuencia de fórmulas mediante la cual se la deduce de los axiomas. Este supuesto no resta generalidad a nuestro argumento, ya que obviamente no cuesta nada darle la forma por él requerida a cualquier deducción de la fórmula numérica $\mathfrak{A}_n$ a partir de los axiomas 1–16. En virtud del mismo, cada fórmula $\mathfrak{A}_k \in \mathcal{L}$ que no sea la última sirve directamente para justificar una única fórmula $\mathfrak{A}_h \in \mathcal{L}$ ($k < h \leq n$) que llamaré la *sucesora* de $\mathfrak{A}_k$ en $\mathcal{L}$. Evidentemente, para que $\mathfrak{A}_h$ sea la sucesora de $\mathfrak{A}_k$ tiene que cumplirse una de las dos condiciones siguientes: o bien (i) $\mathfrak{A}_h$ se deduce de $\mathfrak{A}_k$ por sustitución; o bien (ii) hay un índice $j$ tal que $k \neq j$, $k < h$, $j < h$ y $\mathfrak{A}_h$ se deduce de $\mathfrak{A}_k$ y $\mathfrak{A}_j$ por *modus ponens* (obviamente esto sólo puede ocurrir si $\mathfrak{A}_j = \mathfrak{A}_k \rightarrow \mathfrak{A}_h$ o $\mathfrak{A}_k = \mathfrak{A}_j \rightarrow \mathfrak{A}_h$).

[2] Ackermann expresa el esquema de recursión sólo para los casos $n = 1$ y $n = 2$, pero a la luz de ellos no cabe duda de que la regla general en que está pensando es la formulada arriba.

Queremos demostrar que $\mathfrak{A}_n$ no puede ser la fórmula ¬(0 = 0). Para ello, como se explicó arriba, vamos a convertir a $\mathscr{L}$ en una lista de fórmulas numéricas. Empecemos eliminando las variables. Sea $\mathfrak{A}_p$ la última fórmula de $\mathscr{L}$ que contiene variables. Como $\mathfrak{A}_n$ es una fórmula numérica, es claro que $p < n$ y que $\mathfrak{A}_p$ tiene una sucesora sin variables. Si la sucesora de $\mathfrak{A}_p$ se deduce de ésta por sustitución, efectúo la misma sustitución en $\mathfrak{A}_p$ (en otras palabras, pongo en lugar de $\mathfrak{A}_p$ una copia de su sucesora). Si la sucesora de $\mathfrak{A}_p$ no se deduce de ésta por sustitución, se trata de una fórmula $\mathfrak{C}$ que se deduce por *modus ponens* de premisas $\mathfrak{B}$ y $\mathfrak{B} \rightarrow \mathfrak{C}$, y $\mathfrak{A}_p = \mathfrak{B}$ o $\mathfrak{A}_p = \mathfrak{B} \rightarrow \mathfrak{C}$. Como $\mathfrak{C}$ no contiene variables, todas las variables de $\mathfrak{A}_p$ están contenidas en la subfórmula $\mathfrak{B}$. Las sustituimos simultáneamente en *ambas* premisas. Como figuran en $\mathscr{L}$ por última vez en $\mathfrak{A}_p$, se las puede reemplazar por expresiones más o menos arbitrarias: cada variable individual se reemplaza por 0, cada variable proposicional $A$ se reemplaza por 0 = 0, cada variable funcional $f(a)$ se reemplaza por $a$, cada variable funcional binaria $g(a,b)$ por $a + b$, etc. y todas las variables individuales así introducidas se reemplazan luego por 0 del modo antedicho.[3] Al término de estas operaciones, $\mathfrak{A}_p$ habrá sido sustituida por una fórmula sin variables $\mathfrak{A}'_p$ y $\mathscr{L}$ por una lista $\mathscr{L}'$ cuya última fórmula sin variables es $\mathfrak{A}_q$, con $q < p$. Repitiendo el mismo procedimiento un número finito de veces se formará una lista de fórmulas que no contienen variables. Para convertirla en una lista de fórmulas numéricas hay que eliminar todas las constantes funcionales diferentes de +. Como cada constante funcional φ tiene que haberse definido recursivamente según el esquema antedicho, toda subfórmula sin variables φ( . . . ) que comience con φ se puede traducir mediante un número finito de transformaciones en una fórmula numérica $\mathfrak{z}$, tal que φ( . . . ) = $\mathfrak{z}$ es una fórmula correcta. Sustituimos φ( . . . ) por $\mathfrak{z}$. Como $\mathscr{L}$ es una lista finita de fórmulas finitas, este procedimiento se completa en un número finito de pasos, al cabo de los cuales habremos obtenido una lista $\mathscr{L}^*$ de fórmulas numéricas la última de las cuales es $\mathfrak{A}_n$. Cada fórmula de $\mathscr{L}^*$ o bien (i) es una copia de una fórmula precedente, o bien (ii) se deduce de dos fórmulas precedentes por *modus ponens*, o bien (iii) se obtiene directamente por sustitución de variables en alguno de

3 Aunque Ackermann no lo dice expresamente, tiene que subentender que 0 + 0 se debe reemplazar por 0. De otro modo, no se obtendría a fin de cuentas una lista de *fórmulas numéricas* en el sentido preciso del término.

los axiomas 1–16. En el caso (i), la fórmula en cuestión será correcta, a menos que la precedente que copia sea incorrecta. Es claro asímismo que la conclusión deducida por *modus ponens* de dos premisas correctas es siempre correcta.[4] Por lo tanto, sólo puede haber fórmulas incorrectas en $\mathscr{L}^*$ si alguna de las correspondientes al caso (iii) es incorrecta. Pero eso no es posible: las fórmulas del tipo (iii) obtenidas a partir de uno de los axiomas 1–12 son siempre correctas, no importa que las fórmulas numéricas que reemplacen a cada una de las variables $A$, $B$ y $C$ sean correctas o no;[5] y la sustitución de variables según las reglas arriba descritas en los axiomas 13–16 sólo puede generar fórmulas correctas. Por lo tanto, $\mathfrak{A}_n$, la última fórmula de $\mathscr{L}^*$ y de $\mathscr{L}$, no puede ser la fórmula incorrecta $\neg(0 = 0)$.

Antes de seguir adelante, conviene observar que el sistema de la matemática finita propuesto por Ackermann, con sus 16 axiomas, sus dos reglas de inferencia y su esquema de recursión formaliza justamente la matemática finita que Hilbert y sus discípulos sostienen que debe darse por descontada. Es claro que la prueba de consistencia ofrecida apela a evidencias y emplea modos de razonar propios de esa matemática finita, de modo que en efecto supone aquello mismo que pretende establecer. Sirve, pues, a lo sumo, para verificar que la formalización es adecuada, que no hemos tenido un traspiés al escribir los axiomas con que intentamos representar aquello que está fuera de duda. Además, por cierto, prepara al lector para la prueba siguiente, que emplea los mismos recursos para demostrar la consistencia de un sistema esencialmente más fuerte.

Éste comprende, además de todo lo incluido en el sistema anterior, los "axiomas transfinitos" que enuncio a continuación:[6]

4 Por definición, las fórmulas $\mathfrak{A}$ y $\mathfrak{A} \to \mathfrak{B}$ no pueden ser ambas correctas, a menos que la fórmula $\mathfrak{B}$ lo sea.

5 El lector para quien esto no sea inmediatamente obvio debe comprobarlo construyendo tablas de verdad para esos doce axiomas.

6 Ackermann asigna a los axiomas transfinitos los números 1 al 4; sin embargo, me ha parecido más razonable numerarlos del 17 al 20, ya que los axiomas 1–16 del sistema antes descrito pertenecen también al sistema ahora considerado. Para facilitar la referencia he distinguido con subíndices las diversas formas de los axiomas 17–19. Otros dos puntos en que mi presentación difiere de la original son los siguientes: expreso la negación mediante el signo ¬, en vez de la tilde y el signo ≠ utilizados por Ackermann, y sólo empleo paréntesis redondos, aunque Ackermann también utiliza llaves y corchetes, esto es, los pares { } y [ ]. Por otra parte, respeto todos los pares de paréntesis

$17_0$. $$A(a) \to A(\varepsilon_a A(a))$$
$17_1$. $$A_a f(a) \to A_a((\varepsilon_f A_b f(b))(a))$$

$18_0$. $$A(\varepsilon_a A(a)) \to \pi_a A(a) = 0$$
$18_1$. $$A_a((\varepsilon_f A_b f(b))(a)) \to \pi_f(A_a f(a)) = 0$$

$19_0$. $$\neg A(\varepsilon_a A(a)) \to \pi_a A(a) = 1$$
$19_1$. $$\neg A_a((\varepsilon_f A_b f(b))(a)) \to \pi_f(A_a f(a)) = 1$$

20. $$\neg(\varepsilon_a A(a) = 0)$$

Como se explicó al comienzo, el subíndice en una expresión como $A_a f(a)$ indica que la variable proposicional *A* depende de la variable funcional *f*, no de la variable individual *a*; la presencia de esta última sirve sólo para indicar que la variable *f* es 1-aria. Los Axiomas $17_1$, $18_1$ y $19_1$, que envuelven variables funcionales 1-arias, son propiamente los miembros iniciales de tres secuencias infinitas de axiomas análogos que envuelven variables funcionales *n*-arias, para cada entero positivo *n*. Por ejemplo, el Axioma $17_3$ se enunciaría así:

$17_3$. $$A_{abc} f(a,b,c) \to A_{pqr}((\varepsilon_f A_{pqr} f(p,q,r))(a,b,c))$$

Los operadores ε y π que figuran en los axiomas transfinitos son una ingeniosa creación de Hilbert mediante la cual se pueden definir los signos de cuantificación ∀ y ∃.[7] Ackermann subraya que "ε y π son aquí meros signos

correlativos colocados por él, aunque me parece superfluo el tercer par en el Axioma $17_1$, así como el segundo en los Axiomas $18_1$ y $19_1$.

7 Las siguientes definiciones resumen las ofrecidas por Ackermann (1925, p. 32):

$$\exists u A(u) \leftrightarrow A(\varepsilon_u A(u))$$
$$\forall u A(u) \leftrightarrow A(\varepsilon_u \neg A(u))$$

(aquí, *u* es una variable individual o funcional). Basándose en ellas, no es difícil deducir de los axiomas 1–20 los siguientes teoremas (1925, pp. 32–36):

$$\forall u A(u) \to A(u)$$
$$A(u) \to \exists u A(u)$$
$$\neg \forall u A(u) \leftrightarrow \exists u \neg A(u)$$
$$\forall u \neg A(u) \leftrightarrow \neg \exists u A(u)$$

con los cuales se opera conforme a ciertas reglas" (1925, p. 8). Hay que considerarlos, pues, como caracterizados a cabalidad por los axiomas mismos. Por otra parte, la pretensión de que los Axiomas 1–20 —cuando se los suplementa con el procedimiento para la definición de funciones que describiré luego— bastan para deducir por sustitución y *modus ponens* toda la matemática clásica resulta injustificable e incluso ridícula mientras no se explique cuáles son los ingredientes del pensamiento matemático ordinario que se quiere representar —o reemplazar— con ε y π. Basándome en lo que dice Ackermann, me parece justo explicarlos así: ε y π simbolizan funciones que asignan un número fijo al predicado escrito a su derecha. El Axioma $17_0$ requiere que, si hay números a los que se aplica el predicado *A*, el número designado por $\varepsilon_a A(a)$ sea uno de ellos. Como dice Ackermann, "$\varepsilon_a A(a)$ designa un número del cual es seguro que satisface la aseveración $A(a)$, si es que hay algo que la satisfaga" (p. 8). De otro modo, $\varepsilon_a A(a)$ puede ser cualquier número (aunque es recomendable, a diferencia de Ackermann, asignarle un valor determinado de una vez por todas, por ejemplo, $\varepsilon_a A(a) = 0$). El Axioma $17_1$ impone una exigencia similar a la función designada por $\varepsilon_f A_b f(b)$ en caso que *A* sea un predicado de funciones. Conforme al Axioma 20, el predicado *A* no se aplica en ningún caso al predecesor inmediato del número designado por $\varepsilon_a A(a)$. Conforme a los Axiomas 18 y 19, π simboliza una función que asigna a cada predicado el valor 0 si hay algún número (o una función) a que dicho predicado se aplique, y el valor 1 si no lo hay.[8]

La presencia de ε y π amplia decisivamente el repertorio de funcionales con los que se puede reemplazar una variable de acuerdo con la regla de sustitución. Inicialmente Ackermann parece haber contemplado su utilización irrestricta, pues en el texto de su trabajo dice simplemente que "en lugar de una variable funcional $f(a)$ con un solo hueco, es lícito ahora poner el funcional $\varepsilon_b \mathfrak{A}(a,b)$ o también $\varepsilon_b \mathfrak{A}(a,b,c)$, de modo, pues, que por sustitución en una fórmula $\mathfrak{C}_a f(a)$ se pueden formar las fórmulas nuevas $\mathfrak{C}_a[\varepsilon_b \mathfrak{A}(a,b)]$ o $\mathfrak{C}_a[\varepsilon_b \mathfrak{A}(a,b,c)]$" (p. 9). Pero en una nota agregada "durante la corrección de pruebas", Ackermann prescribe la restricción siguiente:

8 Si hay un número *a* tal que $A(a)$, entonces *A* se aplica al número $\varepsilon_a A(a)$ (Ax. $17_0$) y por lo tanto, $\pi_a A(a) = 0$ (Ax. $18_0$). Si no hay ningún número *a* tal que $A(a)$, tenemos que $\neg A(\varepsilon_a A(a))$, cualquiera que sea el número $\varepsilon_a A(a)$, de suerte que $\pi_a A(a) = 1$ (Ax. $19_0$). El lector podrá extender estas consideraciones al caso —gobernado por los Axiomas $18_n$ y $19_n$ ($n \geq 1$)— en que *A* es un predicado de funciones *n*-arias.

> No está permitido reemplazar una variable funcional $f(a)$ con un funcional $\mathfrak{a}(a)$ en que $a$ figura dentro del alcance (*Bereich*) de un $\varepsilon_f$. Por ejemplo, no se puede poner $(\varepsilon_f\mathfrak{A}(a,f))(0)$ en vez de $f(a)$, pero sí se puede poner $(\varepsilon_f\mathfrak{B}(f))(a) + \varepsilon_b\mathfrak{D}(a,b)$.
>
> (Ackermann 1925, p. 9n.)

Como advierte von Neumann (1927, p. 44), esta restricción es indispensable para que la prueba de consistencia ofrecida por Ackermann sea válida, pero limita severamente el alcance de la misma: la matemática vindicada por ella viene a ser sólo un fragmento de la matemática clásica, equivalente al defendido por Weyl en *El continuo* (1918), o a la teoría ramificada de los tipos *sin* Axioma de Reducibilidad.

También se amplía el esquema para la definición recursiva de nuevas funciones. Ahora se admiten como parámetros no sólo las variables individuales (como $b_1, \ldots, b_{n-1}$ en el esquema R), sino también variables funcionales, que representen a cualquier función previamente caracterizada, inclusive $\varepsilon$ y $\pi$. En vez que expresar esta idea mediante un complejo esquema general, prefiero ilustrarla con tres ejemplos tomados de Ackermann.

$$\begin{aligned} \varphi_b(0,f(b)) &= \mathfrak{a}_b(f(b)) \\ \varphi_b(a+1,f(b)) &= \mathfrak{b}_b(a,f(b),\varphi_c(a,f(c))) \end{aligned} \tag{R1}$$

$$\begin{aligned} \varphi_{bf}(0,g_c(b,f(c))) &= \mathfrak{a}_{bf}(g_c(b,f(c))) \\ \varphi_{bf}(a+1,g_c(b,f(c))) &= \mathfrak{b}_{bf}(a,g_c(b,f(c)),\varphi_{dh}(a,g_d(d,h(c)))) \end{aligned} \tag{R2}$$

$$\begin{aligned} \varphi(0) &= (\varepsilon_f\mathfrak{A}_b(f(b)))(0) \\ \varphi_b(a+1) &= \varphi(a) + (\varepsilon_f\mathfrak{A}_b(f(b)))(a+1) \end{aligned} \tag{R3}$$

La prueba de consistencia del sistema completo sigue esencialmente el método ya descrito, pero, a causa de las novedades introducidas, es muchísimo más complicada. Sea $\mathscr{L} = \mathfrak{A}_1, \ldots, \mathfrak{A}_n$ una lista de fórmulas que constituye una demostración (por sustitución y *modus ponens*) de la fórmula numérica $\mathfrak{A}_n$ a partir de los axiomas 1–20. Supongamos que ya se han eliminado de $\mathscr{L}$

todas las variables libres. Una fórmula de $\mathscr{L}$ derivada del Axioma 17 por sustitución podría entonces ajustarse al patrón siguiente:

$$\mathfrak{A}(\mathfrak{b}) \to \mathfrak{A}(\varepsilon_a\mathfrak{A}(a))$$

donde $\mathfrak{b}$ es un funcional determinado y $\mathfrak{A}(a)$ es una fórmula cuya única variable libre es $a$. Naturalmente, tanto $\mathfrak{b}$ como $\mathfrak{A}(a)$ pueden contener los símbolos ε y π. Por ejemplo, $\mathfrak{b}$ podría ser un funcional de la forma

$$\varepsilon_f\mathfrak{B}_{ei}(f(e),\varepsilon_c\mathfrak{C}(c),\varepsilon_d\mathfrak{D}(f(i),d)))(0)$$

y $\mathfrak{A}(a)$ la fórmula $\varepsilon_b(a = b) = 0$. Entonces, la formula en cuestión se representaría más adecuadamente así:

$$\varepsilon_b(\varepsilon_f\mathfrak{B}_{ei}(f(e),\varepsilon_c\mathfrak{C}(c),\varepsilon_d\mathfrak{D}(f(i),d)))(0) = b) = 0 \to \varepsilon_b(\varepsilon_a(\varepsilon_c(a = c) = 0) = b) = 0$$

donde $\mathfrak{B}$, $\mathfrak{C}$ y $\mathfrak{D}$ representan fórmulas cuya complejidad está apenas insinuada. Para reducir $\mathscr{L}$ a una lista $\mathscr{L}$* de puras fórmulas numéricas hay que eliminar todas las constantes funcionales definidas recursivamente y también ε y π. En los casos más sencillos, se puede proceder por "ensayo y error". Consideremos una vez más el patrón $\mathfrak{A}(\mathfrak{b}) \to \mathfrak{A}(\varepsilon_a\mathfrak{A}(a))$, mas suponiendo ahora que ε no figura en $\mathfrak{A}(a)$ y que el funcional $\mathfrak{b}$ es una fórmula numérica. En tal caso, nuestro procedimiento habitual redundará en la sustitución de $\mathfrak{b}$ por cierto numeral $\mathfrak{z}$. Tentativamente, reemplazamos $\varepsilon_a\mathfrak{A}(a)$ por 0. La fórmula resultante $\mathfrak{A}(\mathfrak{z}) \to \mathfrak{A}(0)$ es correcta a menos que $\mathfrak{A}(\mathfrak{z})$ sea correcto y $\mathfrak{A}(0)$ incorrecto, y entonces el propio $\mathfrak{z}$ es un sustituto adecuado para $\varepsilon_a\mathfrak{A}(a)$.

Pero el caso general es muchísimo más complicado. Como subraya Ackermann, con sólo mirar una fórmula no se puede saber de antemano cuántas sustituciones de ε y π habrá que efectuar para que estos símbolos desaparezcan del todo. Considérese, por ejemplo, una expresión como ésta:

$$\varepsilon_a\mathfrak{A}(\varphi_b(\varepsilon_d\mathfrak{B}(d),\varepsilon_c\mathfrak{K}(c,b)),a))$$

donde la función φ se definerecursivamente así:

$$\varphi_b(0, f(b)) = f(0)$$

$$\varphi_b(a + 1, f(b)) = \varphi_b(a, f(b)) + f(a + 1)$$

Si $\varepsilon_d\mathfrak{B}(d)$ se deja reemplazar lícitamente por cierto numeral $\mathfrak{z}$, la expresión considerada se transforma en $\varepsilon_a\mathfrak{A}(\varphi_b(\mathfrak{z},\varepsilon_c\mathfrak{K}(c,b)),a))$, la cual, mediante la definición de $\phi$, se reduce a:

$$\varepsilon_a\mathfrak{A}(\varepsilon_c\mathfrak{K}(c,0) + \varepsilon_c\mathfrak{K}(c,1) + \ldots + \varepsilon_c\mathfrak{K}(c,\mathfrak{z}),a)$$

El número de símbolos ε que hay que sustituir en esta nueva expresión depende obviamente del número de unos en el numeral $\mathfrak{z}$. Esta dificultad es, por cierto, una consecuencia de la inclusión de variables funcionales como parámetros en el esquema de recursión y afecta también, por lo tanto, a cualquier constante funcional definida según ese esquema. Si la constante φ entra en su propia definición recursiva, cuando la reemplacemos con ayuda de su definición en una fórmula dada, puede que acabe figurando en la fórmula resultante un mayor número de veces que en la propia fórmula inicial.

Para establecer que la reiterada sustitución de las funciones que figuran en un funcional cualquiera lo transforma en una fórmula numérica al cabo de un número finito de operaciones, Ackermann concibe un modo de asignarle a cada funcional un índice característico (formado por varios números), define un orden lineal entre esos índices y demuestra que el índice de un funcional dado disminuye con cada sustitución prescrita hasta reducirse, tras un número finito de ellas, al índice mínimo, característico de las fórmulas numéricas. La demostración es tortuosa y no voy a reproducirla aquí (en la nota siguiente explico como se construye el índice de un funcional en el caso más simple en que éste no contiene los signos ε y π).[9] Como ya señalé, su

9 Para determinar el índice característico de un dado funcional $\mathfrak{S}$, en el cual no figuran los signos ε y π, Ackermann tiene en cuenta el orden en que se han definido recursivamente las funciones que figuran en $\mathfrak{S}$. Obsérvese que este orden no depende del funcional en cuestión sino del modo cómo se ha construido el sistema. En todo caso, la primera de esas funciones es la adición. Diré que la función ψ es *mayor o igual* que la función φ si ψ = φ o si φ precede a ψ en dicho orden. Si φ y ψ figuran en el funcional $\mathfrak{S}$, diré que ψ *está subordinada* a φ (en $\mathfrak{S}$) si ψ figura dentro del paréntesis gobernado por φ y dentro del alcance de ψ hay una variable ligada por un subíndice de φ. Si χ está subordinada a ψ y ψ está subordinada a φ, decimos que χ también está subordinada a φ. Ackermann asigna a cada funcional un *rango* con respecto a cada una de las funciones que figuran en él. El rango del funcional $\mathfrak{S}$ con respecto a una

alcance está severamente limitado por la restricción que Ackermann tuvo que imponer a su regla de sustitución.

---

cierta función φ depende de lo que llamaré el *nivel* relativo a φ de las funciones mayores o iguales que φ que figuran en 𝔖. Sea Φ el conjunto de estas funciones. Si ψ ∈ Φ y ninguna función de Φ está subordinada a ψ en 𝔖, el nivel de ψ relativo a φ es 1. Si ψ ∈ Φ y $n \geq 1$ es el más alto nivel relativo a φ que alcanza una función de Φ subordinada a ψ (en 𝔖), el nivel de ψ relativo a φ es $n + 1$. El rango de 𝔖 con respecto a φ es el más alto nivel relativo a φ que tiene en 𝔖 una función mayor o igual que φ. Si φ no figura en 𝔖 diré que el rango de 𝔖 respecto de φ es 0 (Ackermann no menciona esta posibilidad, pero hay que tenerla en cuenta para que sea verdad, como él dice, que cada funcional tiene un rango con respecto a cada función recursiva). Cada funcional 𝔖 contiene una colección finita de funcionales que llamaré sus *componentes*. Por ejemplo, los componentes de $\psi_d(\psi_e(0 + 1 + 1,\varphi_b(e,b)),\varphi_b(d,b))$ son: $0 + 1 + 1$, $e$, $b$, $d$, $\varphi_b(e,b)$, $\psi_e(0 + 1 + 1,\varphi_b(e,b))$, $\varphi_b(d,b)$ y $\psi_d(\psi_e(0 + 1 + 1,\varphi_b(e,b)),\varphi_b(d,b))$. Cada componente $\mathfrak{S}_k$ de un dado funcional 𝔖 tiene un rango determinado con respecto a cada una de las funciones que se han definido recursivamente en el sistema; la lista de estos números, dispuesta de modo que las funciones respectivas queden ordenadas de mayor a menor, constituye la *combinación de rangos* de $\mathfrak{S}_k$. Las combinaciones de rangos de los funcionales contenidos en 𝔖 se ordenan de mayor a menor según la convención siguiente: $\langle a_1, \ldots ,a_r\rangle > \langle b_1, \ldots ,b_s\rangle$ si $r > s$, o si $r = s$ y hay un $k$ ($1 \leq k \leq r$) tal que $a_k < b_k$ pero $a_j = b_j$ si $j < k$ (nótese que en virtud de las convenciones adoptadas cada combinación de rangos contiene el mismo número de enteros no negativos, a saber, el número de las funciones que se han definido recursivamente en el sistema). A cada combinación presente en la lista le asignamos el número total de veces que figuran en 𝔖 funcionales que posean esa combinación. Además, asignamos el 0 a cada posible combinación de rangos ausente de la lista que sea menor que la mayor de las presentes. La lista de números así generada, dispuesta en el orden de las correspondientes combinaciones de rangos, es el *indice* del funcional 𝔖. Los índices de distintos funcionales se ordenan de mayor a menor según la misma convención que aplicamos a las combinaciones de rango. Si el funcional 𝔖 forma parte de una fórmula numérica tiene que ser un numeral, esto es, un 0, o un 0 seguido de una o más copias de la configuración '+ 1'. Según nuestras estipulaciones, si 𝔖 es un numeral, él mismo es su solo componente y su rango es 1 respecto de la adición y 0 respecto de todas las otras funciones del sistema. Por lo tanto, su combinación de rangos es la mínima, $\langle 0, \ldots ,0,1\rangle$, y su índice es el menor de todos, $\langle 1\rangle$. En cambio, si 𝔖 no es un numeral, su combinación de rangos tendrá al menos un entero positivo en una posición distinta de la última y su índice, por ende, constará de dos o más números. En las pp. 16–18, Ackermann (1925) prescribe sustituciones aplicables a cualquier funcional 𝔖 que no sea un numeral y no contenga variables libres ni los signos ε y π, las cuales demostrablemente generan un funcional de índice más bajo que aquél al que se aplican y lo transforman al cabo de un número finito de pasos en un funcional de índice $\langle 1\rangle$. En el resto de su trabajo, Ackermann extiende estos procedimientos a funcionales que contengan una o más copias de ε y π.

### *2.6.2 Von Neumann (1927)*

El escrito de von Neumann “Sobre la teoría hilbertiana de la prueba” (1927) trata el mismo tema que Ackermann (1925), pero es mucho más luminoso y fácil de leer. El autor formula y clasifica los axiomas de que supuestamente puede deducirse toda la matemática clásica de un modo que deja bien claros los límites de la prueba de consistencia ofrecida. Ésta no cubre cabalmente el sistema propuesto, sino sólo el fragmento que depende de cinco de sus seis grupos de axiomas, aunque von Neumann conjetura —con demasiado optimismo— que la consistencia de *todos* los axiomas podrá establecerse con el mismo método que aplica con éxito a ese fragmento. Dicho método tiene especial interés para nosotros, porque anticipa los utilizados después de 1930 en el contexto de la llamada teoría de modelos (que estudiaremos en la Parte 3). Antes de explicarlo, me referiré a ciertas observaciones de carácter general que von Neumann enuncia en las primeras secciones de su escrito.

Primero, sobre la “matemática clásica”. Von Neumann entiende por esto “la matemática [ . . . ] como era universalmente reconocida (*allgemein anerkannt*) hasta que surgieron los críticos de la teoría de conjuntos: le pertenecen esencialmente todos los métodos de la teoría de conjuntos, pero no la teoría de conjuntos abstracta; y, por de pronto, no conoce cardinalidades mayores que la del continuo” (p. 256). No sabría identificar con precisión ese cuerpo universalmente reconocido de métodos y teoremas matemáticos. Me siento inclinado a creer que para von Neumann la “matemática clásica” comprendía, por ejemplo, los recursos que empleó Lebesgue en su teoría de la integración; pero inmediatamente después de la caracterización citada nos advierte que no resolverá, por el momento, si ella incluye o no el Axioma de Selección.

Para probar la consistencia de la matemática clásica hay que formalizar con absoluto rigor “todo su aparato de aseveraciones y demostraciones”. Esto incluye, por cierto, una formalización de la aritmética elemental o teoría de los números naturales. Con todo, “no es posible edificar la teoria de la prueba a menos que ya esté disponible a priori el número entero positivo con todas esas propiedades suyas que se pueden derivar a la manera intuicionista, es decir, sustantivamente (*inhaltlich*)” (p. 258). En otras palabras, la aritmética finita *à la* Skolem (Capítulo 2.5), aunque derivable en la matemática formalizada, tiene que darse por descontada al emprender la formalización. Ocu-

rre, pues, que los mismos conceptos tienen dos modos de manifestarse, (i) dentro del formalismo que se investiga y (ii) en la reflexión "metateórica" acerca de él. Además de los enteros positivos, "todas las relaciones lógicas" aparecen bajo esta doble faz (p. 258). Von Neumann reconoce que esto puede causar confusión (p. 258).

Von Neumann distingue cinco clases de símbolos simples: variables (constituidas por la letra $x$ seguida de un subíndice numérico), constantes ($C$ con subíndice), operaciones ($O$ con subíndice y exponente), abstracciones ($A$ con subíndice), y signos de puntuación (coma, paréntesis). Operaciones son los conectivos lógicos, la identidad, la función 'el siguiente de…', la suma, la multiplicación, etc.; el exponente indica "el número de conceptos a los que se aplica una dada operación" (p. 259), mientras el subíndice sirve para distinguir diversas operaciones con el mismo exponente (aunque en la práctica von Neumann las representa con los signos habituales, por ejemplo, $\rightarrow$ en vez de $O_1{}^2$, = en vez de $O_2{}^2$, etc.). Abstracciones son los operadores que ligan variables, esto es, los signos de cuantificación universal y existencial (que representaré, como de costumbre, con $\forall$ y $\exists$) y el operador $\tau$ (afín al operador $\varepsilon$ que estudiamos en la Sección 2.6.1). Von Neumann da una definición recursiva precisa de lo que se llama una fórmula. Ella permite decidir en todos los casos si una combinación cualquiera de símbolos simples es o no una *fórmula*. Dada una fórmula, hay siempre un y sólo un modo de construirla a partir de una determinada colección finita de símbolos simples, mediante un número finito de aplicaciones sucesivas de las reglas que forman la definición de fórmula. Este principio es familiar a cualquiera que haya estudiado lógica moderna y está claramente ilustrado en nuestro Apéndice IX. Von Neumann subraya su "fundamental importancia": "un formalismo en que no se cumpla sería rechazado por todos como ininteligible e inservible" (p. 261). Una fórmula sin variables libres es una fórmula *normal*. En la teoría de la prueba se trata siempre de fórmulas normales (aunque, como es obvio, dentro de una fórmula normal $\varphi$ puede figurar una subfórmula $\psi$ que contenga variables libres ligadas por abstracciones que preceden a $\psi$ en $\varphi$).

En la presentación de von Neumann, la pieza clave de una teoría formalizada es una regla $\mathfrak{R}$ para construir cierto género de fórmulas normales que llamaremos *axiomas*. $\mathfrak{R}$ tiene que concebirse de modo que permita decidir, en presencia de cualquier fórmula normal, si ésta es o no un axioma. $\mathfrak{R}$ se expresa mediante una serie de esquemas. Un *axioma* es cualquier fórmula

obtenida reemplazando por fórmulas normales (o, en algunos casos, por variables) ciertos elementos de uno de esos esquemas. Por ejemplo, cualquier fórmula obtenida reemplazando α y β por fórmulas normales en el esquema $(\alpha \rightarrow (\beta \rightarrow \alpha))$ es un axioma según la regla $\mathfrak{MR}$ característica de la matemática clásica. En contraste con los sistemas de Frege, Whitehead-Russell y Ackermann, tal reemplazo no es aquí una operación del sistema formal — una deducción conforme a una "regla de sustitución"— sino un procedimiento metateórico, un paso más en la *formulación* —intrínsecamente interminable— del propio sistema. Von Neumann advierte que ese procedimiento "no debe confundirse en ningún caso con las operaciones formalistas, pues es un proceso sustantivo (*inhaltlicher*)": el signo de sustitución —$\beta_{\zeta/\xi}$ en nuestro Apéndice IX (p. 485), $\mathrm{Subst}\binom{\xi}{\zeta}\beta$ en el escrito de von Neumann— "significa" algo (1927, p. 264).

Relativamente a una determinada regla de construcción de axiomas $\mathfrak{R}$, una *fórmula demostrable* se define así:

[$D_1$] Todo axioma es una fórmula demostrable.

[$D_2$] Si ψ es una fórmula normal y φ y $(\varphi \rightarrow \psi)$ son fórmulas ya reconocidas como demostrables, entonces ψ es una fórmula demostrable.

Como von Neumann expresamente destaca, esta definición nos enseña a construir (*aufstellen*) fórmulas demostrables, mas no a decidir (*entscheiden*) frente a una fórmula normal dada, si ella es demostrable o no. En esto contrasta con la definición recursiva de 'fórmula' que aparece en nuestro Apéndice IX (similar a la de von Neumann). Esto se debe a que, para decidir que una cierta fila de signos es una *fórmula* basta examinar los signos de que consta y el modo cómo se combinan, sin que haya que tener en cuenta a ningún objeto que no esté dado con ella. En cambio, para determinar que una cierta fórmula normal ψ que no sea un axioma es *demostrable*, hay que conocer una fórmula demostrable φ tal que $(\varphi \rightarrow \psi)$ también sea demostrable, y obviamente el mero examen de ψ no revela esa fórmula ϕ.

> Parece, pues, que no hay ninguna vía para descubrir el criterio universal de decisión (*allgemeine Entscheidungskriterium*) sobre si una dada fórmula normal *a* es demostrable. Por cierto, actualmente no podemos probar nada

> a este respecto. No hay tampoco ninguna indicación de cómo podría probarse dicha indecidibilidad. Pero esta incertidumbre no nos impide constatar que hoy en día no es posible decidir universalmente si una fórmula normal cualquiera *a* es demostrable o no (relativamente a la regla de construcción de axiomas que se describirá luego). Y que ello sea indecidible es incluso la *conditio sine qua non* para que tenga sentido hacer matemáticas con los métodos heurísticos de hoy. El día mismo que la indecidibilidad cese, también dejará de existir la matemática en el sentido actual; en su lugar habría una receta completamente mecánica con ayuda de la cual cualquiera podría decidir acerca de cualquier aseveración si se la puede o no demostrar.
>
> (von Neumann 1927, p. 266)

En el formalismo de von Neumann, la fórmula construida anteponiendo el operador $O_1{}^1$ a una fórmula cualquiera $\alpha$ se llama la *negación* de $\alpha$. Escribiré $\neg\alpha$ en vez de $O_1{}^1\alpha$ (von Neumann escribe $\sim\alpha$). Una regla de construcción de axiomas $\mathfrak{R}$ se dice *consistente* si no hay ninguna fórmula normal $\alpha$ tal que $\alpha$ y $\neg\alpha$ sean ambas demostrables relativamente a $\mathfrak{R}$. La consistencia de una dada regla $\mathfrak{R}$ quedará establecida sin lugar a dudas si podemos definir una partición de todas las fórmulas normales en dos clases con las propiedades siguientes:

[$W_1$] Dada una fórmula normal a siempre es posible decidir, en un número finito de pasos, a cuál de las dos clases pertenece.

[$W_2$] Ninguna fórmula normal pertenece a la misma clase que su negación.

[$W_3$] Todos los axiomas pertenecen a la misma clase.

[$W_4$] La fórmula normal $(\alpha \to \beta)$ pertenece a la misma clase que los axiomas si y sólo si la fórmula $\beta$ pertenece a esa clase o la fórmula $\alpha$ pertenece a la otra clase.

Una partición que cumpla estos cuatro requisitos es una *valuación* (*Wertung*) de $\mathfrak{R}$. Como es evidente, en una valuación todas las fórmulas demostrables pertenecen a la misma clase que los axiomas. Von Neumann llama a esa clase, la clase *R* (por *richtig*, 'correcto'), pero yo la llamaré *V*. A la otra clase la llamo, con von Neumann, *F* (*falsch*, en alemán, significa a la vez 'incorrecto' y 'falso').

Von Neumann enuncia una regla $\mathfrak{M}\mathfrak{R}$ que según él es capaz de generar todos los axiomas que hacen falta para deducir toda la matemática clásica

(por *modus ponens*; cf. la cláusula [$D_2$] de la definición de 'fórmula demostrable'). Los axiomas generables con arreglo a $\mathfrak{M}\mathfrak{N}$ se clasifican en seis grupos. El Grupo I caracteriza los conectivos lógicos $\neg$ y $\rightarrow$ (formalmente, $O_1^1$ y $O_1^2$) y basta para generar o deducir por *modus ponens* a todas las tautologías. El Grupo II caracteriza la identidad = (formalmente, $O_2^2$). El Grupo III es una versión sin cuantificadores de los cuatro primeros axiomas de Peano: caracteriza la constante 0 (formalmente, $C_1$; léase *cero*) y los operadores $Z$ y +1 (formalmente, $O_2^1$ y $O_3^1$; $Z$ precede y +1 sigue a su argumento; léase *…es un número* y *el siguiente de…*). Los tres primeros grupos equivalen, pues, a los Axiomas 1–16 de Ackermann. El Grupo VI y último es el grupo de las definiciones (para cuya representación formal von Neumann ofrece esquemas). Concentraremos nuestra atención en los Grupos IV y V.

Al Grupo IV pertenece toda fórmula que se obtenga a partir de uno de los siguientes esquemas $IV_1$-$IV_4$ cuando $\beta$ es reemplazada por una fórmula normal, $\xi$ por una variable y $\alpha$ por una fórmula sin otra variable libre que la elegida para reemplazar a $\xi$, o —trivialmente— por una fórmula normal. (Conforme a la convención de nuestro Apéndice IX, $\alpha_{\beta/\xi}$ sería entonces la expresión obtenida al sustituir uniformemente la variable que reemplace a $\xi$ por la formula normal que reemplace a $\beta$ en la fórmula que reemplace a $\alpha$):

$IV_1$. $\forall\xi\alpha \rightarrow \alpha_{\beta/\xi}$
$IV_2$. $\alpha_{\beta/\xi} \rightarrow \exists\xi\alpha$
$IV_3$. $\alpha_{\tau_\xi\alpha/\xi} \rightarrow \forall\xi\alpha$
$IV_4$. $\exists\xi\alpha \rightarrow \alpha_{\tau_\xi\neg\alpha/\xi}$

Siguiendo a von Neumann, el sentido del operador $\tau$ se puede explicar así: una fórmula $\alpha$ cuya única variable libre es $\xi$ representa un predicado que la fórmula normal $\alpha_{\beta/\xi}$ atribuye al objeto designado por la expresión $\beta$; si hay objetos a los que dicho predicado *no* se aplica, $\tau_\xi\alpha$ designa uno de esos objetos; pero si no hay ningún objeto al que no se aplique el predicado referido, $\tau_\xi\alpha$ designa cualquier cosa. (Como se puede ver, $\tau$ es simplemente el operador complementario o "dual" del operador $\varepsilon$ utilizado por Ackermann; cf. la nota 8). Si $\tau$ se entiende de esta manera, los esquemas $IV_3$ y $IV_4$ responden a la interpretación familiar de los símbolos $\rightarrow$, $\forall$ y $\exists$. Tenemos en efecto (i) que el predicado representado por la fórmula $\alpha$ con variable libre $\xi$ se aplica al objeto particular designado por $\tau_\xi\alpha$ sólo si no hay ningún objeto

al que dicho predicado no se aplique (esquema $IV_3$), y (ii) que si hay objetos al que se aplique el predicado representado por $\alpha$, a esos objetos no se aplica el predicado representado por $\neg\alpha$ y $\tau_\xi\neg\alpha$ designa uno de ellos, esto es, un objeto al que se aplica el predicado representado por $\alpha$ (esquema $IV_4$).

Von Neumann observa acertadamente que "los típicos modos de inferencia 'transfinitos' o 'impredicativos'" representados por los esquemas del Grupo IV no bastan para edificar la matemática clásica. Ello requiere además una cierta parte de la teoría de conjuntos.

> Es necesario destacarlo expresamente: Hay dos puntos en que el edificio de la matemática clásica está inseguro y expuesto a los ataques de los escépticos, a saber, el concepto "todos" y el concepto de "conjunto". Estas dos cosas fundamentalmente diferentes no deben identificarse (como suele ocurrir), pero tampoco puede permitirse que una de ellas nos haga olvidar la otra. La crítica de la matemática comenzó por el concepto de "conjunto" y lentamente ha avanzado hasta el de "todos", que hoy, empero, es el principal punto de ataque de los intuicionistas. Pero no hay que olvidar que, aun cuando sus objeciones contra "todos" hayan sido refutadas en cierto sentido, con eso no se ha rescatado aún el concepto de conjunto.
>
> (von Neumann 1927, p. 271)

Conforme al patrón establecido en sus escritos sobre teoría de conjuntos (Sección 1.8.4), von Neumann sustituye el concepto de conjunto por el concepto de función. Introduce un operador binario $\Phi$ (formalmente, $O_2^3$) tal que $\Phi(\xi,\eta)$ designa el valor de la función designada por $\xi$ en el argumento designado por $\eta$. El Grupo V de axiomas incluye todas las fórmulas que se obtengan reemplazando, en el siguiente esquema, $\xi$ y $\eta$ por dos variables distintas y $\alpha$ por una fórmula sin variables libres o cuya única variable libre sea la elegida para reemplazar a $\eta$:

$$\exists\xi\forall\eta(Z\eta \rightarrow (\Phi(\xi,\eta) = \alpha))$$

Von Neumann explica esto así: $\forall\eta(Z\eta \rightarrow (\Phi(f,\eta) = \alpha))$ implica que, cualquiera que sea el objeto $c$, $Zc \rightarrow (\Phi(f,c) = \alpha_{c/\xi})$; en otras palabras, la función $f$ representa la expresión $\alpha_{c/\xi}$ en el dominio de los números. Por lo tanto, lo que el esquema dice es sencillamente que toda expresión está representada por una cierta función en el dominio de los números.

> Y ésta es justamente la pretensión de la teoría de conjuntos: Todas las aseveraciones sobre *x*, por compleja que sea su estructura y por intrincado que sea el modo como *x* figura en ellas, pueden reducirse a la forma normal "*x* elemento de *M*", donde *M* es independiente de *x*. Como en vez de conjuntos consideramos funciones, en lugar de aseveraciones hallamos expresiones, en lugar de la equivalencia lógica la simple identidad y en lugar de la forma normal "*x* elemento de *M*" la forma normal "valor de la función *F* para el argumento *x*".
>
> (von Neumann 1927, p. 273)

Como indiqué al comienzo, von Neumann no logra probar la consistencia de $\mathfrak{MR}$. Sólo prueba la consistencia de $\mathfrak{MR}$(I–IV,VI), la regla para la construcción de los axiomas de todos los grupos excepto el V. Su prueba, como él mismo señala, no va más lejos que la de Ackermann; pero es muchísimo más clara y elegante. Von Neumann prepara el terreno con una prueba simple y contundente de la consistencia de $\mathfrak{MR}$(I–III) (comparable a la prueba de Ackermann concerniente a sus Axiomas 1–16). Ella consiste en definir una valuación de la regla $\mathfrak{MR}$(I–III) por las siguientes estipulaciones. [i] Todas las variables y todas las constantes menos una pertenecen a la clase *V*; la constante exceptuada, que von Neumann escribe informalmente $W_F$, pertenece a la clase *F*. [ii] Si $\phi$ es una fórmula que consta de una operación *n*-aria y *n* fórmulas $\alpha_1, \ldots ,\alpha_n$ cuya clasificación ya se ha determinado, entonces [ii$_1$] si $\phi$ es $\neg\alpha_1$, $\phi$ es *V* si $\alpha$ es *F* y $\phi$ es *F* si $\alpha$ es *V*; [ii$_2$] si $\phi$ es $(\alpha_1 \rightarrow \alpha_2)$, $\phi$ es *V* si y sólo si $\alpha_1$ es *F* o $\alpha_2$ es *V*; [ii$_3$] si $\phi$ es $(\alpha_1 = \alpha_2)$, $\phi$ es *V* si y sólo si $\alpha_1$ y $\alpha_2$ son idénticas; [ii$_4$] si $\phi$ es $Z\alpha_1$, $\phi$ es *V* si y sólo si $\alpha_1$ consta de un 0 seguido de ninguna, una o varias copias de la expresión '+1';[10] [ii$_5$] si $\phi$ no obedece a ninguna de las cuatro cláusulas precedentes, $\phi$ es *V*. [iii] Si $\phi$ es una fórmula que consta de una abstracción seguida de una variable y una fórmula cualquiera, $\phi$ es *V*.[11] Esencialmente la prueba de consistencia

10 Formalmente, esta cláusula se puede expresar con más claridad: [ii$_4$] si $\phi$ es $O_2{}^1\alpha_1$, $\phi$ es *V* si y sólo si $\alpha_1$ es la constante $C_1$, o consta de $C_1$ precedida de una o varias copias de $O_3{}^1$.

11 Las estipulaciones [ii$_5$] y [iii] tienen el efecto de asegurar —trivialmente— la clasificación de fórmulas que carecen de toda importancia en la teoría determinada por la regla $\mathfrak{MR}$(I–III) cuya consistencia se trata de probar. La necesidad de clasificar tales fórmulas desaparece si restringimos el lenguaje de la teoría en cuestión, estipulando que no

consiste en establecer que cualquier axioma de los grupos I–III, esto es, cualquier fórmula obtenida por reemplazo a partir de los esquemas característicos de esos grupos, pertenece a la clase *V*. Una vez que esto ha sido asegurado, la cláusula [$ii_2$] garantiza que cada fórmula demostrable es *V* y la cláusula [$ii_1$] garantiza que la negación de una fórmula demostrable es *F* y por ende no puede ser demostrable.

Von Neumann señala enfáticamente que no sabríamos dar una valuación de $\mathfrak{MR}$, ni siquiera de $\mathfrak{MR}$(I–IV). Toda valuación contiene un procedimiento efectivo para clasificar a cualquier fórmula y por lo tanto permite decidir en un número finito de pasos si una dada fórmula $\alpha$ es irrefutable (si es clasificada como *V*, en cuyo caso $\neg\alpha$ tiene que ser *F*) o indemostrable (si es clasificada como *F*). Aun cuando ello no bastaría para resolver la cuestión arriba planteada de si una dada fórmula es o no demostrable, von Neumann considera "sumamente inverosímil" que exista un procedimiento universal para alcanzar siquiera esta decisión más débil "con respecto a todos los problemas matemáticos posibles" (1927, p. 276). Pero aunque la prueba de consistencia por valuación tenga que confinarse a $\mathfrak{MR}$(I–III), la consistencia de $\mathfrak{MR}$(I–IV,VI) puede probarse mediante lo que von Neumann llama una valuación parcial. Y en 1927 von Neumann no ha perdido la esperanza de extender este método a $\mathfrak{MR}$.[12] No puedo entrar aquí en las complejidades de su aplicación a fórmulas que contengan cuantificadores o el operador $\tau$. Pero daré, para terminar, una caracterización general del mismo.

Sea $\mathfrak{R}$ una regla para la construcción de axiomas. Una valuación parcial de $\mathfrak{R}$ es un procedimiento que define, para cada lista finita $\mathfrak{S}$ de axiomas construidos conforme a $\mathfrak{R}$ una partición de las fórmulas normales en dos clases $V_{\mathfrak{S}}$ y $F_{\mathfrak{S}}$ con las propiedades siguientes:

[$W_1'$] Dada una fórmula normal $\alpha$ siempre es posible decidir, en un número finito de pasos, si pertenece a la clase $V_{\mathfrak{S}}$ o a la clase $F_{\mathfrak{S}}$.

---

contiene abstracciones, ni más operaciones que las mencionadas en las cláusulas [$ii_1$]–[$ii_4$].

12 Von Neumann 1927, p. 276, n. 8 atribuye a Julius König la idea de probar la consistencia de la matemática mediante una valuación y a Hilbert las ideas en que se basa el concepto de valuación parcial.

[$W_2'$] La fórmula normal $\alpha$ pertenece a la clase $V_{\mathfrak{S}}$ si y sólo si $\neg\alpha$ pertenece a la clase $F_{\mathfrak{S}}$.

[$W_3'$] Todos los axiomas de la lista $\mathfrak{S}$ pertenecen a la clase $V_{\mathfrak{S}}$.

[$W_4'$] La fórmula normal $(\alpha \to \beta)$ pertenece a la clase $V_{\mathfrak{S}}$, si y sólo la fórmula $\beta$ pertenece a la clase $V_{\mathfrak{S}}$ o la fórmula $\alpha$ pertenece a la clase $F_{\mathfrak{S}}$.

Como puede verse, toda valuación es a la vez una valuación parcial[13] con la notable propiedad de que la partición $\{V_{\mathfrak{S}}, F_{\mathfrak{S}}\}$ que ella asocia a cada lista finita de axiomas $\mathfrak{S}$ no depende de la índole de $\mathfrak{S}$. Una valuación parcial corriente no sería tan neutra. Sin embargo, basta dar una valuación parcial de $\mathfrak{R}$ para estar seguro de que $\mathfrak{R}$ es consistente. Para demostrarlo, supongamos que hay una valuación parcial de $\mathfrak{R}$ —que llamaré $\mathfrak{W}_{\mathfrak{R}}$— pero que $\mathfrak{R}$ es inconsistente. Esto último implica que hay dos fórmulas normales $\alpha$ y $\neg\alpha$ que son ambas demostrables relativamente a $\mathfrak{R}$. Ahora bien, a la luz de las cláusulas [$D_1$] y [$D_2$] de la definición de fórmula demostrable, si $\alpha$ es demostrable relativamente a $\mathfrak{R}$ tiene que haber una lista finita $\mathfrak{L}_\alpha$ de axiomas construidos con arreglo a $\mathfrak{R}$, tal que, o bien $\mathfrak{L}_\alpha$ tiene un solo miembro y ese miembro es $\alpha$, o bien $\alpha$ se deduce de los miembros de $\mathfrak{L}_\alpha$ mediante un número finito de aplicaciones de *modus ponens*. Otro tanto vale para $\neg\alpha$ y una lista finita $\mathfrak{L}_{\neg\alpha}$. Sea $\mathfrak{L}$ la unión de las listas $\mathfrak{L}_\alpha$ y $\mathfrak{L}_{\neg\alpha}$ y $\{V_{\mathfrak{L}}, F_{\mathfrak{L}}\}$ la partición de las fórmulas normales asociada a $\mathfrak{L}$ por la valuación parcial $\mathfrak{W}_{\mathfrak{R}}$. Entonces, por [$W_3'$] y [$W_4'$], $\alpha$ y $\neg\alpha$ pertenecen a la clase $V_{\mathfrak{L}}$. Pero esto es imposible, por [$W_2'$]. Por lo tanto, si hay una valuación parcial de $\mathfrak{R}$, $\mathfrak{R}$ no puede ser inconsistente.

[13] Puesto que, para cada entero positivo $k \leq 4$, [$W_k$] implica [$W_k'$].

### *2.6.3 Herbrand (1931)*

En una nota sobre la “Consistencia (*non-contradiction*) de los axiomas aritméticos” (1929), comunicada a la Academia de París, Jacques Herbrand deslinda con luminosa precisión la teoría axiomática cuya consistencia habría logrado establecer von Neumann (1927) y esboza un método suyo, más sencillo, para llegar al mismo resultado. Herbrand distingue dos teorías, que llamaré $T_1$ y $T_2$ (él las llama Th. 1 y Th. 2). Ambas se refieren a objetos (*individus*) llamados números (*nombres*), entre los cuales hay uno designado por la constante ‘0’. Hay un único predicado ‘=’ (que se escribe entre los dos términos a que se aplica) y una “función descriptiva” que asigna a cada número $a$ el número designado por ‘$a + 1$’. $T_1$ tiene tres grupos de axiomas:

[I] $a = a$, $a = b \rightarrow b = a$, $a = b \wedge b = c \rightarrow a = c$,

$$a + 1 = b + 1 \leftrightarrow a = b \qquad \neg(a + 1 = 0)$$

[**II**] Todas las fórmulas de la secuencia

$$\neg(a = a + 1), \neg(a = a + 1 + 1), \neg(a = a + 1 + 1 + 1), \ldots$$

[**III**] Todas las fórmulas que se obtengan reemplazando una letra dada por una expresión cualquiera (formada con los medios descritos) en los axiomas anteriores o también —como debemos suponer, aunque Herbrand no lo diga expresamente— en una tautología[14] formada con letras y los cinco conectivos $\neg$, $\vee$, $\wedge$, $\rightarrow$ y $\leftrightarrow$.

La única regla de inferencia de $T_1$ es *modus ponens*. La teoría $T_2$ tiene los mismos axiomas, pero admite el uso de variables ligadas por cuantificadores (“variables aparentes”, como dice Herbrand). Hemos de suponer que admite, además de *modus ponens*, las reglas para la introducción y manipulación de tales variables descritas en lo que Herbrand (1929) llama “nuestra Nota precedente” (esto es, Herbrand 1928): reglas de generalización universal y

[14] Esto es, en palabras de Herbrand, una “proposición cuyo valor veritativo (*valeur logique*) es la verdad, cualquiera que sea el valor veritativo de las letras que figuran en ella” (Herbrand, EL, p. 21).

existencial y las reglas familiares para pasar[15] de $\neg\forall x\phi x$ a $\exists x\neg\phi x$, de $\neg\exists x\phi x$ a $\forall x\neg\phi x$, de $p \vee \forall x\phi x$ a $\forall x(p \vee \phi x)$, y de $p \vee \exists x\phi x$ a $\exists x(p \vee \phi x)$, o viceversa. Herbrand observa que la consistencia de $T_1$ es fácil de probar. Luego señala que la teoría cuya consistencia von Neumann (1927) demuestra "por un método bastante complicado" es precisamente $T_2$. Bosqueja "los fundamentos de una demostración más simple" y concluye:

> Para completar la aritmética sería menester introducir las definiciones recursivas y el axioma de inducción total, cuya consistencia nuestro método —igual que el de von Neumann— no permite demostrar. Hay que advertir, sin embargo, que si se introduce solamente el axioma de inducción total, sin definiciones recursivas, éste resulta ser una consecuencia de los axiomas precedentes y, por lo tanto, no es contradictorio.
>
> (Herbrand, EL, p. 26)

La demostración bosquejada en la nota de 1929 se presenta en detalle en la tesis doctoral de Herbrand (1930). Su simplicidad se debe justamente a que utiliza el principal resultado de esa tesis, el poderoso Teorema de Herbrand, cuya demostración, eso sí, es muy complicada. En virtud del mismo, una fórmula $\Phi$ de $T_2$ es demostrable (con los recursos señalados) sólo si ciertas fórmulas sin cuantificadores, construibles a partir de $\Phi$ según un procedimiento que Herbrand especifica, cumplen ciertas condiciones efectivamente verificables. Como veremos en el Capítulo 2.7, esas condiciones son tales que si una fórmula $\Phi$ de $T_2$ se halla en el caso descrito es evidentemente imposible que su negación $\neg\Phi$ se halle también en el mismo caso.

Herbrand trata nuevamente el tema de manera más concisa, pero en cierto modo más completa y eficaz, en el artículo póstumo "Sobre la consistencia de la aritmética" (1931) que terminó en Göttingen el 14 de julio de 1931, trece días antes de caer a un precipicio en el macizo de Pelvoux (tenía veintitrés años). Allí vindica la observación final arriba transcrita, mediante el invento genial que examinaremos ahora. Herbrand considera ahora una teo-

15 "Règles de passage" es como Herbrand llama (en otros escritos) a las próximas cuatro reglas y sus inversas. Enuncia también otras, que envuelven los conectivos $\wedge$ y $\rightarrow$. El deseo de ser breve explica —pienso— la omisión de estas últimas en Herbrand 1928, a pesar de que en ese escrito se acepta como primitivo el conectivo $\rightarrow$ (escrito $\supset$, *à la Russell*), además de $\neg$ y $\vee$.

ría que llamaré $T_3$, formalizada como $T_1$ y $T_2$ en un cálculo predicativo de primer orden con sólo una constante (0), un predicado diádico (=) y un functor primitivo (+1). $T_3$ retiene los axiomas del grupo I, llamado aquí Grupo A, pero en lugar del grupo II se introduce el siguiente Grupo B (esquema de "inducción total"):

[B] Si $\Phi(x)$ representa una fórmula que contiene la variable libre $x$, entonces la fórmula construida según el esquema siguiente es un axioma:

$$(\Phi(0) \wedge \forall x(\Phi(x) \rightarrow \Phi(x + 1))) \rightarrow \forall x \Phi(x)$$

Un tercer grupo de axiomas [C] consiste en las ecuaciones requeridas para definir un número arbitario pero finito de functores $k$-arios (para uno o más enteros positivos $k$). Con respecto a esas ecuaciones, Herbrand sólo estipula que (i) no pueden contener variables ligadas, (ii) permiten "calcular efectivamente" el valor de la función respectiva para cualquier lista idónea de números ($k$ números cualesquiera, si el functor es $k$-ario), y (iii) se puede "demostrar de modo intuicionista"[16] que el valor así calculable está bien determinado en cada caso.[17]

El último grupo de axiomas de $T_3$ autoriza la deducción de generalizaciones mediante la siguiente estipulación:

[16] Cf. la definición de esta expresión por Herbrand citada en la p. 125, nota 15. Designa precisamente los métodos sustantivos (*inhaltlich*), finitistas, de la metamatemática hilbertiana.

[17] Herbrand no define aquí el concepto de *cálculo efectivo*, pero da dos ejemplos: (1) Si $\alpha$ y $\beta$ son, respectivamente, un término y un functor 1-ario previamente definidos, puede introducirse un nuevo functor $f$ caracterizado por las ecuaciones $f(0) = \alpha$ y $f(x + 1) = \beta(fx)$. (2) Si los términos $a$ y $b$ y las funciones $\langle x,y \rangle \mapsto x + y$ y $x \mapsto x - 1$ han sido previamente definidos, se puede introducir un nuevo functor 3-ario $\varphi$ caracterizado por las ecuaciones $\varphi(0,a,b) = a + b$, $\varphi(x,a,1) = a$ y $\varphi(x + 1,a,b) = \varphi(x,a,\varphi(x + 1,a,b - 1))$. En el Capítulo 2.11 nos ocuparemos con este tema del cálculo efectivo de funciones numéricas. Como allí veremos, la primera definición precisa del concepto (Gödel 1934), equivalente a todas las que se han propuesto más tarde, se inspiró en una comunicación privada de Herbrand a Gödel.

**[D]** Si $\Phi(x)$ representa una fórmula sin variables ligadas, tal que, cuando se la considera sustantivamente (esto es, como enunciando una propiedad de los enteros y no como un mero símbolo), se puede probar mediante "procedimientos intuicionistas" que $\Phi(x)$ es verdadera para todo $x$, entonces $\forall x\Phi(x)$ es un axioma.

La teoría $T_3$ así caracterizada es por lo menos tan fuerte como la aritmética clásica. Herbrand no pretende haber demostrado su consistencia y reconoce que, a la luz del teorema recién publicado por Gödel (1931), bien puede ser imposible demostrarla.[18] Su prueba de consistencia vindica sólo ciertos *fragmentos de la aritmética*, más amplios en todo caso que el vindicado por Ackermann (1925) y von Neumann (1927). Sus resultados pueden formularse así:

⟨1⟩ La teoría determinada por los Grupos A, B, C y D es consistente si las fórmulas $\Phi(x)$ a que se refiere el Grupo B no contienen variables ligadas.

⟨2⟩ La teoría determinada por los Grupos A, B, C y D es consistente aunque las fórmulas $\Phi(x)$ a que se refiere el Grupo B contengan variables ligadas, si dichas fórmulas no contienen ningún functor de los definidos en el Grupo C (en otras palabras, si el único functor que figura en ellas es el functor 1-ario '+1' caracterizado por los dos últimos axiomas del Grupo A).

Para lograr estos resultados, Herbrand se vale de la ocurrencia que arriba llamé genial y que ahora paso a explicar. Consiste en definir, para cada fór-

[18] Entiéndase, con los métodos que él llama intuicionistas (veáse arriba, nota 16, y Capítulo 2.1, nota 14). La siguiente observación con que termina Herbrand (1931) se entenderá mejor una vez que hayamos estudiado el descubrimiento de Gödel en el Capítulo 2.10:

> No es imposible que todo razonamiento intuicionista pueda representarse (*se faire*) en una aritmética con los axiomas A y B, sin admitir en C más que la adición y la multiplicación ordinarias. De ser así, aun la consistencia de la aritmética ordinaria sería indemostrable.
>
> (Herbrand, EL, p. 232)

mula $\Phi(x)$ a que se refiere el Grupo B (sin restricciones) una función $\boldsymbol{\varepsilon}$ con la siguiente propiedad: si $a$ es el número más pequeño tal que $\Phi(a)$ es falsa, $\boldsymbol{\varepsilon}(x) = 0$ cuando $x < a$ y $\boldsymbol{\varepsilon}(x) = a$ cuando $x \geq a$. La definición rigurosa de $\boldsymbol{\varepsilon}$ está dada por el siguiente Grupo E de esquemas (el lector debe verificar que la función $\boldsymbol{\varepsilon}$ aquí definida tiene en efecto la propiedad indicada):

[E1] $$\boldsymbol{\varepsilon}(0) = 0$$

[E2] $$(\Phi(0) \wedge \neg\Phi(x + 1) \wedge \boldsymbol{\varepsilon}(x) = 0) \to \boldsymbol{\varepsilon}(x + 1) = x + 1$$

[E3] $$\neg(\Phi(0) \wedge \neg\Phi(x + 1) \wedge \boldsymbol{\varepsilon}(x) = 0) \to \boldsymbol{\varepsilon}(x + 1) = \boldsymbol{\varepsilon}(x)$$

[E4] $$\boldsymbol{\varepsilon}(x) = y + 1 \to \boldsymbol{\varepsilon}(y + 1) = y + 1 \wedge \boldsymbol{\varepsilon}(y) = 0$$

Mediante un razonamiento tan simple como brillante, Herbrand demuestra que el esquema de inducción total B es una consecuencia de E1–E4. Supongamos, en efecto, que cierta fórmula $\Phi(x)$ satisface las condiciones E1–E4, pero que el respectivo esquema B es falso. Esto último significa suponer que $\Phi(0)$ y $\forall x(\Phi(x) \to \Phi(x + 1))$ son verdaderas, pero que hay un número $x$ tal que $\neg\Phi(x + 1)$. En tal caso, E2 y E3 implican que $\boldsymbol{\varepsilon}(x + 1) = x + 1$ si $\boldsymbol{\varepsilon}(x) = 0$ y $\boldsymbol{\varepsilon}(x + 1) = \boldsymbol{\varepsilon}(x)$ si $\boldsymbol{\varepsilon}(x) \neq 0$. Así pues, en cualquier caso, $\boldsymbol{\varepsilon}(x + 1) \neq 0$, de modo que hay un número $y$ tal que $\boldsymbol{\varepsilon}(x + 1) = y + 1$. Por lo tanto, en virtud de E4, $\boldsymbol{\varepsilon}(y + 1) = y + 1$ y $\boldsymbol{\varepsilon}(y) = 0$. Esto implica que $\boldsymbol{\varepsilon}(y + 1) \neq \boldsymbol{\varepsilon}(y)$. De esta desigualdad y E3 se infiere (por *modus tollens*) que $(\Phi(0) \wedge \neg\Phi(y + 1) \wedge \boldsymbol{\varepsilon}(y) = 0)$ y por ende que $\neg\Phi(y + 1)$. Supusimos inicialmente que $\forall x(\Phi(x) \to \Phi(x + 1))$ es verdadera, lo cual implica que $\forall x(\neg\Phi(x + 1) \to \neg\Phi(x))$. Concluimos, por lo tanto, que $\neg\Phi(y)$, de modo que $y \neq 0$. Hay, pues, un número $z$ tal que $y = z + 1$. Tenemos que $\boldsymbol{\varepsilon}(z + 1) = \boldsymbol{\varepsilon}(y) = 0 \neq z + 1$. De esta desigualdad y E2 se infiere (por *modus tollens*) que $\neg(\Phi(0) \wedge \neg\Phi(z + 1) \wedge \boldsymbol{\varepsilon}(z) = 0)$, lo cual, combinado con E3, permite concluir (por *modus ponens*) que $\boldsymbol{\varepsilon}(z) = \boldsymbol{\varepsilon}(z + 1) = \boldsymbol{\varepsilon}(y) = 0$. Por otra parte, como $\Phi(0)$ y $\neg\Phi(z + 1)$ —o sea, $\neg\Phi(y)$— son verdaderas, $\neg(\Phi(0) \wedge \neg\Phi(z + 1) \wedge \boldsymbol{\varepsilon}(z) = 0)$ implica que $\boldsymbol{\varepsilon}(z) \neq 0$. ¡Contradicción! Es claro, entonces, que el Grupo E puede sustituir al Grupo B.

Si las fórmulas $\Phi(x)$ a que se refieren los Grupos B y E no contienen variables ligadas, los esquemas E1–E4 caen según Herbrand dentro del Grupo C, “pues en ese caso se puede constatar efectivamente si $\Phi(x)$ es o no

verdadera y las hipótesis E permiten calcular efectivamente los valores $\boldsymbol{\varepsilon}(x)$ de modo unívoco y no contradictorio" (EL, p. 228). Para establecer el resultado ⟨1⟩ basta entonces probar la consistencia de los grupos A, C y D. Como se verá en el próximo capítulo, ello no es difícil, una vez se dispone del Teorema de Herbrand. También el resultado ⟨2⟩ es una consecuencia fácil de ese teorema.

## 2.7 EL *ENTSCHEIDUNGSPROBLEM* Y EL TEOREMA DE HERBRAND

El programa de Hilbert busca vindicar la matemática clásica con métodos finitistas o "intuicionistas"[1] a través del estudio de su formulación en una escritura conceptual o cálculo lógico, como el **CP1=** de nuestro Apéndice IX, o el que utilizan Whitehead y Russell. Tales cálculos tienen una propiedad que resulta ser importantísima para ese propósito: para decidir si un objeto cualquiera es un signo del cálculo, si una fila de tales signos es una fórmula, o si una lista de fórmulas es una prueba (de la última fórmula de la lista a partir de una o más de las anteriores) basta efectuar un número finito de operaciones simples, cada una de las cuales arroja un resultado inequívoco e indubitable. Por ejemplo, para decidir que & no es un signo de nuestro **CP1=** basta compararlo con cada uno de los 10 signos del mismo; para establecer que una lista $\langle\alpha_1, \ldots, \alpha_n\rangle$ de fórmulas de la teoría $T_1$ de Herbrand 1929 (Sección 2.6.3) es una prueba de $\alpha_n$ a partir de los axiomas [**I**], [**II**] y [**III**] de $T_1$ basta determinar, con respecto a cada fórmula $\alpha_k$ $(1 \leq k \leq n)$ de la lista, que $\alpha_k$ es un axioma o es idéntica a la subfórmula $\gamma$ de una fórmula $\alpha_h = (\beta \rightarrow \gamma)$, donde $h < k$ y $\beta = \alpha_j$ para algún índice $j < k$. La índole intuitiva y rutinaria de tales métodos de decisión recuerda los familiares algoritmos para calcular la suma, el producto, el mínimo común múltiplo, etc., de una lista de enteros. Como veremos en el Capítulo 2.11, aquí no hay sólo un parecido, sino una afinidad esencial. Usaré por eso el término genérico 'algoritmo' para referirme a cualquier procedimiento como esos. La noción es algo nebulosa, pero en el Capítulo 2.11 veremos que hay varios modos (equivalentes) de hacerla precisa. Diré que un cálculo lógico K es *efectivo* si hay un algoritmo para decidir si una lista de objetos es una prueba en K (lo cual supone, claro está, que haya algoritmos para decidir si un objeto cualquiera es un signo, una fórmula o un axioma de K).

[1] Sobre el uso hilbertiano del adjetivo 'intuicionista', *vide supra*, p. 125, nota 15.

Porque el cálculo de Whitehead y Russell es efectivo, sus propiedades se pueden estudiar "metamatemáticamente" con métodos "intuicionistas". Como las teorías fundamentales de la matemática clásica se dejan representar en él, Hilbert concibió la esperanza de demostrar la consistencia de esas teorías a través de la investigación metamatemática de su representación en el cálculo. Bastaba establecer que una fórmula dada —por ejemplo, la que representa la desigualdad $0 \neq 0$— no puede ser la última fórmula de una prueba a partir de los axiomas de la teoría investigada. Obviamente, el problema quedaría resuelto si hubiera un algoritmo para decidir, ante una fórmula cualquiera, si ella es o no deducible de los axiomas. La efectividad del cálculo no garantiza que tal algoritmo exista. Si el cálculo es efectivo es posible, claro está, determinar en un número finito de pasos si cierto despliegue de signos que termina con la fórmula en cuestión obedece o no a las reglas que que gobiernan las pruebas. Pero otra cosa muy distinta es decidir con medios finitos si en el conjunto infinito de las pruebas posibles hay o no una que termine con esa fórmula.

En los años veinte empezó a usarse el vocablo alemán *Entscheidungsproblem* ('problema de la decisión') para designar, por antonomasia, este último problema. Formulémoslo con un poco más de precisión. Sea *T* una teoría matemática representada en un cálculo efectivo K. Sea $\mathscr{A}$ la clase de fórmulas de K que son axiomas de *T*. Sea $\alpha$ una fórmula de K. Decimos que $\alpha$ es deducible de $\mathscr{A}$ en K —simbólicamente, $\mathscr{A} \vdash_K \alpha$— si hay una lista $\mathscr{L}$ de fórmulas de K tal que [i] $\alpha$ es la última fórmula de $\mathscr{L}$ y [ii] si $\beta$ es cualquier fórmula de $\mathscr{L}$, o bien $\beta \in \mathscr{A}$, o bien $\beta$ se infiere conforme a una de las reglas de inferencia de K de fórmulas que preceden a $\beta$ en $\mathscr{L}$. Un procedimiento de decisión (*Entscheidungsverfahren*) para *T* es un algoritmo que permita decidir, en el caso de cualquier fórmula $\alpha$ de K, si $\mathscr{A} \vdash_K \alpha$ o no. El *Entscheidungsproblem* de la teoría *T* consiste en hallar un procedimiento de decisión para *T*. Como luego explicaré, Post (1921) resolvió positivamente el *Entscheidungsproblem* para el llamado cálculo proposicional, que es un pequeño fragmento del sistema de *Principia Mathematica*. Animados en parte por este éxito y también, sin duda, por la índole puramente combinatoria de la construcción de pruebas en un cálculo efectivo, varios seguidores de Hilbert creyeron que podrían hallar un procedimiento de decisión para el fragmento mayor (correspondiente a nuestro **CP1=**) que juzgaban suficiente para representar la matemática clásica. Vimos en la Sección 2.6.2 que von Neumann

(1927, p. 266), con característica sagacidad, expresaba dudas con respecto a la viabilidad de este proyecto, y no hay indicios de que Hilbert mismo se haya interesado por él.[2] En cambio, para el joven Herbrand ese era "el problema fundamental de la lógica matemática" (EL, pp. 64, 170), "en cierto modo, el problema más general de las matemáticas" (EL, p. 210).[3] Si no hubiera creído que podía resolverlo, difícilmente habría llevado adelante la endiablada demostración del teorema que lleva su nombre y que todavía sirve de base a la solución positiva del *Entscheidungsproblem* en los casos especiales en que ésta es posible (cf. Dreben y Goldfarb 1979).

Antes de seguir adelante, debo señalar que la definición de 'procedimiento de decisión' que he dado aquí, ajustándome a las formulaciones de

[2] Antes bien, en su importante artículo "Sobre lo infinito", Hilbert da por descontado que no puede haber un método general —no digamos un algoritmo— para resolver todos los problemas matemáticos. Tras reiterar su convicción de que en la matemática no hay problemas insolubles, agrega: "Por cierto, mi teoría de la prueba no puede indicar en general una vía por la cual todo problema matemático pueda resolverse: tal vía tampoco existe" (1925, p. 180).

[3] La última frase está tomada del resumen de la tesis doctoral de Herbrand, redactado por él mismo y publicado sin firma en los Anales de la Universidad de París en 1931. Vale la pena leer el párrafo completo que la precede:

> Russell mostró en *Principia Mathematica* que para hacer matemáticas se puede usar, en vez del lenguaje ordinario, una suerte de estenografía, de lenguaje simbólico que utiliza sólo un número muy limitado de signos [ . . . ], cuyas combinaciones forman las frases. Pero fue más lejos y eso es lo que nos importa: mostró que todas las demostraciones matemáticas posibles dependen de unas cuantas reglas simples de razonamiento que él enuncia; que en toda teoría matemática determinada se comienza por admitir como verdaderas ciertas proposiciones —los axiomas de esa teoría— que se traducen en combinaciones de signos; que todas las demostraciones que se pueden hacer en la teoría se reducen al empleo sucesivo de ciertas reglas bien determinadas que permiten fabricar nuevas proposiciones verdaderas a partir de proposiciones ya reconocidas como verdaderas, y que, si se traduce toda la demostración al lenguaje simbólico, estas reglas pueden enunciarse como reglas para determinadas combinaciones de los signos de este lenguaje. Así vemos, pues, que el problema "¿Puede tal proposición demostrarse en una teoría que tiene tales axiomas?" es un problema concerniente a los signos de este lenguaje y sus combinaciones, y que es susceptible de un tratamiento matemático. Esa es la forma más general del problema que los alemanes llaman el "Entscheidungsproblem". Es, en cierto modo, el problema más general de las matemáticas.
>
> (Herbrand, EL, pp. 209–10)

Herbrand (y a las demandas del programa de Hilbert), difiere de la que doy en el Apéndice IX. Según ésta, un procedimiento de decisión para un cálculo lógico es un algoritmo para determinar la *validez* —no la *deducibilidad*— de sus fórmulas. Como ambos conceptos se entrelazan en la historia que voy a relatar, conviene distinguirlos mediante prefijos: '*v*-' para validez, '*d*-' para deducibilidad.[4] A un hilbertiano de estricta observancia tiene que importarle la *d*-decisión, que se refiere a una propiedad sintáctica de las fórmulas, independiente de su posible significado, y no la *v*-decisión, que se refiere a una cuestión semántica, concerniente al significado y la verdad.[5] Sin embargo, los dos trabajos más importantes dedicados al *Entscheidungsproblem* en la revista de Hilbert (Behmann 1922, Bernays y Schönfinkel 1928) lo caracterizan exclusivamente en términos de *v*-decisión. Aún más llamativo es el

4 Más adelante, encontraremos un tercer concepto de decisión, en el título y en el texto del célebre artículo de Gödel, "Sobre proposiciones formalmente indecidibles de *Principia Mathematica* y sistemas afines" (1931). Sea K un cálculo lógico efectivo con un signo de negación que representaré con $\neg$. Sea $R$ una relación $n$-ádica entre números naturales ($n \geq 1$). $R$ es decidible en K en el sentido de Gödel 1931 si K tiene una fórmula con $n$ variables libres $\phi(\xi_1, \ldots, \xi_n)$, tal que $\varnothing \vdash_K \phi(\alpha_1/\xi_1, \ldots, \alpha_n/\xi_n)$ si $\alpha_1, \ldots, \alpha_n$ son los nombres en K de una lista de números que tiene la relación $R$ y $\varnothing \vdash_K \neg\phi(\alpha_1/\xi_1, \ldots, \alpha_n/\xi_n)$ si $\alpha_1, \ldots, \alpha_n$ son los nombres en K de una lista de números que no tiene la relación $R$. (La notación $\phi(\alpha_1/\xi_1, \ldots, \alpha_n/\xi_n)$ se explica en el Apéndice IX, p. 485). Conviene advertir que Gödel no es enteramente responsable de esta confusión terminológica. En la definición que acabo de parafrasear, donde —como todos los traductores— digo 'decidible', el original no dice 'entscheidbar', sino 'entscheidungsdefinit', que propiamente significa 'determinado con respecto a la decisión'. Como Gödel estaba indudablemente familiarizado con los conceptos de *v*-decisión y *d*-decisión arriba explicados y con el llamado *Entscheidungsproblem*, cabe pensar que con la palabra 'entscheidungsdefinit' no quiso decir 'decidible' —en una nueva y caprichosa acepción— sino 'determinado *con vistas a* (o *para los efectos de*) la *d*-decisión'. Pero Gödel usa invariablemente 'unentscheidbar' ('indecidible') para calificar las relaciones que no son *entscheidungsdefinit* —así como las proposiciones que las aseveran— hasta en el mismo título de su trabajo.

5 Como incisivamente advierte Kleene (1952, p. 175), "la noción misma de validez no es finitista (*finitary*) en el caso de un dominio infinito y una fórmula que contenga [un predicado $n$-ádico]. Pues supone que el valor de una función sea [0] para todas las funciones lógicas de $n$ variables consideradas como valores de ese predicado, y la clase de tales funciones lógicas es indenumerable y por ende sólo es concebible (como solemos pensar) en términos del infinito actual". (Los corchetes marcan dos pequeños cambios que he introducido en el texto de Kleene para ajustarlo a nuestro Apéndice IX).

hecho de que tanto Post como Herbrand, que ponen sus miras resueltamente en la *d*-decisión, la abordan, como veremos, con lo que en el fondo son métodos de *v*-decisión, que sirven al interés declaradamente sintáctico de estos autores porque ocurre que se ocupan con cálculos en los que toda fórmula válida es a la vez deducible.

El concepto semántico de *fórmula válida* definido en el Apendice IX (para el **CP1**= y el **CP2**=) es, por cierto, posterior a estos trabajos: lo introdujo Tarski en un escrito que apareció en polaco en 1933, en alemán sólo en 1935 (Capítulo 3.1). Pero un concepto análogo está implícito ya en los argumentos con que Frege justifica los axiomas y reglas de inferencia de su cálculo BS (1879) y figura destacadamente en la obra de Schröder (VAL, 1890ss.), quien, inspirándose en Boole y sobre todo en Peirce, fundó en Alemania otra tradición de lógica matemática.[6] Esa tradición culmina en el articulo de Leopold Löwenheim, "Sobre posibilidades en el cálculo de relativos" (1915). Dicho artículo, que por su simbolismo obsoleto nos resulta hoy casi ilegible, fue estudiado acuciosamente por Skolem, Herbrand y Gödel. Contiene la primera versión del célebre Teorema de Löwenheim y Skolem. Rompiendo el orden cronológico, dejo su estudio para el Capítulo 3.2, porque la obra de Tarski a que se refiere el Capítulo 3.1 nos ayudará a entender mejor el punto de vista semántico presupuesto en la concepción misma del teorema. Pero debo dar aquí algunas indicaciones sobre dicho punto de vista, pues Gödel lo adopta en las investigaciones que estudiaremos en los Capítulos 2.8 y 2.10, y el propio Herbrand, cuya ortodoxia hilbertiana le impedía adoptarlo, lo explota como un recurso metódico.

Para no fatigar al lector con el aprendizaje del simbolismo de Löwenheim, evitaré citar sus fórmulas.[7] Éstas son ecuaciones, casi siempre de la forma $\Phi = 0$ ó $\Phi = 1$, donde $\Phi$ es una expresión que contiene variables, que pueden ser de distintos tipos. En las ecuaciones de primer orden[8] —las únicas a que

6 En cambio, en *Principia Mathematica* el concepto de validez brilla por su ausencia.

7 La nota introductoria a la traducción inglesa de Löwenheim 1915 en van Heijenoort 1967 contiene indicaciones sobre el "cálculo de relativos" de Peirce-Schröder que bastan para descifrar ese escrito y los escritos tempranos de Skolem.

8 Löwenheim las llama "ecuaciones numéricas" (*Zahlgleichungen*). El nombre que les doy corresponde al utilizado en la traducción inglesa de Löwenheim 1915 citada en la nota precedente. Se justifica porque toda ecuación numérica en el sentido de Löwenheim puede expresarse mediante una fórmula de nuestro cálculo predicativo de primer orden

voy a referirme aquí— las variables son de un solo tipo y recorren un mismo dominio de objetos (*Denkbereich*, “universo del discurso”), no vacío pero por lo demás indeterminado, que Löwenheim, siguiendo a Schröder, llama $1^1$. Dependiendo de su estructura, Φ denota una clase de *n*-tuplos de objetos de $1^1$ o expresa una aseveración sobre tales objetos o clases. El significado del signo ‘1’ a la derecha de la ecuación depende de la índole de la expresión Φ a la izquierda: si ésta nombra una clase de *n*-tuplos, 1 es la clase de todos los *n*-tuplos posibles; pero si Φ es una aseveración, ‘Φ = 1’ expresa que Φ es verdadera. 0 es en todo caso el complemento booleano del 1 pertinente: la clase vacía, si 1 es una clase; “lo falso”, si 1 es una aseveración. Como es obvio, una ecuación que en cierto dominio de objetos se cumple (vale, está satisfecha) como quiera que se fijen los valores de sus variables, puede no cumplirse en otro dominio. Una ecuación satisfecha en todo caso, cualquiera que sea el dominio $1^1$, es una *ecuación idéntica*. Löwenheim llama *Fluchtgleichung* —digamos *ecuación huidiza*— a una ecuación que *no es* idéntica, pero que se cumple en cada dominio *finito*, como quiera que se fijen los valores de sus variables; y llama *Haltgleichung* —diré *ecuación detenible*— a una ecuación que falla para alguna asignación de valores a sus variables en un dominio finito.[9] Evidentemente, la clasificación de las ecuaciones en idénticas, huidizas y detenibles es una partición.

El Teorema 2 de Löwenheim (1915) dice que toda ecuación huidiza de primer orden falla para alguna asignación de valores a sus variables en un dominio enumerable. Este es el enunciado original del célebre teorema generalizado más tarde por Skolem.[10] Pero en el presente contexto nos interesa

---

**CP1=** y toda fórmula del **CP1=** puede representarse mediante una ecuación numérica en el sentido de Löwenheim.

9 La idea que inspira esta curiosa terminología me parece ser la siguiente. Podemos tratar de refutar la validez universal de una ecuación examinando si se cumple en dominios de 1, 2, 3, . . . objetos. Si ella es una *Haltgleichung* este proceso se detiene con algún entero positivo *n*; pero en el caso de una *Fluchtgleichung* el proceso continúa indefinidamente a pesar de que la ecuación no es universalmente válida. Así, la ecuación huidiza es una que logra escapar a la refutación que sin embargo merece.

10 La formulación habitual del Teorema “ascendente” de Löwenheim-Skolem generaliza el siguiente corolario, derivado por contraposición del enunciado anterior: si una ecuación de primer orden está satisfecha —como quiera que se fijen los valores de sus variables— en un dominio enumerable y en cada dominio finito, esa ecuación es una ecuación idéntica.

más el Teorema 4. Digamos que una ecuación de primer orden es unaria si contiene expresiones que denotan clases de objetos de $1^1$, pero no contiene ninguna expresión que denote, para algún $n > 1$, una clase de $n$-tuplos de tales objetos. Tales ecuaciones pueden evidentemente representarse en aquel fragmento del **CP1=** que no contiene predicados poliádicos. Este fragmento del **CP1=** (o de otro cálculo lógico equivalente o análogo) suele llamarse cálculo de predicados monádicos. El Teorema 4 dice que ninguna ecuación unaria es huidiza. Esto significa que toda ecuación unaria es detenible o idéntica. La demostración —que Löwenheim solamente bosqueja— provee un método constructivo que permite establecer, con respecto a cada ecuación unaria en que figuran $k$ predicados monádicos diferentes, o bien que ella es idéntica, o bien que no está satisfecha para alguna asignación de valores en un dominio de $2^k$ objetos. Tal método constituye en efecto un procedimiento de *v*-decisión para el cálculo de predicados monádicos. El Teorema 4 de Löwenheim (1915) constituye pues la primera solución positiva parcial —*avant la lettre*— del *Entscheidungsproblem* (en sentido semántico, es decir, como problema de la *v*-decisión).[11]

La tesis doctoral de Emil Post, "Introducción a una teoría general de las proposiciones elementales" (1921) plantea explícitamente y resuelve el problema de la *d*-decisión con respecto a una parte del sistema deductivo de *Principia Mathematica*. Dicha parte comprende precisamente lo que en la jerga de nuestro Apéndice IX llamaríamos fórmulas verifuncionales y los axiomas que consisten en fórmulas de esa clase. Su nombre tradicional es *cálculo proposicional*, porque se entiende que los predicados 0-ádicos ("variables proposicionales") que figuran en tales fórmulas representan proposiciones indeterminadas.[12] Siguiendo a Whitehead y Russell (PM, ***1**), Post

[11] Así lo entiende Behmann en su escrito "Contribuciones al álgebra de la lógica, especialmente al *Entscheidungsproblem*" (1922). La principal contribución que ahí se ofrece es precisamente el Teorema 4 de Löwenheim 1915, reformulado como solución al problema de la *v*-decisión del cálculo de predicados monádicos. Behmann da una demostración detallada y relativamente clara, basada en la de Löwenheim; pero aún más clara es la elegante demostración de Bernays y Schönfinkel (1927) que presento en el Apéndice XIV.

[12] El rechazo filosófico de la idea misma de proposición inspiró luego el nombre *sentential calculus* que suele, curiosamente, vertirse al castellano como 'cálculo sentencial', quizás porque la palabra 'oración' —equivalente exacto de la inglesa 'sentence'— tiene, además de su significado gramatical, uno religioso (la alergia a la connotación paralela

llama 'proposiciones elementales' a las fórmulas de este cálculo, pero como aquí hemos reservado el adjetivo 'elemental' para sus ingredientes inanalizables (que otros llaman 'fórmulas atómicas'), las llamaré, como de costumbre, 'fórmulas verifuncionales' o simplemente 'fórmulas' (si no hay peligro de confusión). Post adopta un punto de vista decididamente sintáctico: estudia el cálculo proposicional de *Principia* y ciertas generalizaciones del mismo exclusivamente como objetos formales ("as purely *formal developments*"—p. 266).[13] Sus resultados "*conciernen* a la lógica de las proposiciones pero *no* están *incluidos* en ella" (p. 265). Esta observación anticipa el distingo hilbertiano entre los razonamientos *formales,* que la metamatemática estudia, y los razonamientos *sustantivos* (*inhaltlich*), que ella practica; pero, en contraste con Hilbert, que quería rebatir el intuicionismo con los limitados recursos que éste admite, Post está dispuesto a aprovechar "cualquier instrumento de la lógica y la matemática que [encuentre] útil" (p. 266).[14]

La descripción del cálculo proposicional, organizado como sistema deductivo, está contenida en cuatro postulados. El Postulado I da la regla para construir nuevas fórmulas a partir de una o dos fórmulas dadas. Los Postulados II y III dan reglas para obtener nuevos asertos a partir de uno o dos asertos dados. El Postulado IV despliega los asertos primitivos. Conviene advertir que las mismas letras minúsculas *p*, *q*, *r*, con o sin subíndices numéricos, que Whitehead y Russell usaban como variables del cálculo proposicional (el equivalente de nuestros predicados 0-ádicos), Post las utiliza como parte del idioma de trabajo (o sea, como "letras esquemáticas" o "variables metalingüísticas") para representar variables proposicionales cualesquiera, y también, aparentemente (en el Postulado I), para representar fórmulas cualesquiera. Por otra parte, usa bien decididamente las mayúscu-

---

habría operado, quizás, en la otra dirección si entre los profesores de filosofía de habla castellana fuese mayor el número de ex-penados y no tan grande el de ex-seminaristas).

13 Para "una formulación general de este punto de vista", Post remite al capítulo VI de la obra de C. I. Lewis, *Survey of Symbolic Logic* (1918), desgraciadamente omitido en la reimpresión de 1960. Leemos allí que "un sistema matemático es cualquier conjunto de filas de signos reconocibles (*strings of recognizable marks*) en el cual algunas filas se adoptan inicialmente y las demás se derivan de ellas mediante operaciones ejecutadas conforme a reglas que son independientes de cualquier significado asignado a las signos" (p. 355; en el original, este pasaje está destacado en cursiva).

14 Como veremos en el Capítulo 2.8, Gödel 1930 adoptará una postura metódica similar.

las $P$ y Q para representar fórmulas indeterminadas (en el Postulado III). Como es habitual, Post usa los conectivos del cálculo para nombrarlos al discurrir sobre éste.[15] El cálculo proposicional de *Principia Mathematica* tiene dos conectivos primitivos, $\sim$ y $\vee$, en términos de los cuales se definen estos tres: $\supset$, $\cdot$ y $\equiv$. Post usa estos cinco signos como en ese libro, pero aquí los reemplazaré por $\neg$, $\vee$, $\rightarrow$, $\wedge$ y $\leftrightarrow$, respectivamente. Por lo demás, usaré libremente la jerga de nuestro Apéndice IX. Para la puntuación usaré paréntesis en vez de los sistemas de puntos que Post imita de *Principia*. Post se vale del signo fregeano $\vdash$ prefijado a un esquema de fórmulas para indicar que cualquier fórmula construida según ese esquema es una aseveración (*assertion*). Entiendo que este signo no pertenece al cálculo, sino al discurso sobre el cálculo.

A continuación enuncio los postulados con que Post caracteriza el cálculo proposicional de *Principia Mathematica*.

I Si $p$ y $q$ son fórmulas verifuncionales (*elementary propositions*), $\neg p$ y $(p \vee q)$ son fórmulas verifuncionales.[16]

II Si una aseveración $t$ contiene una variable del cálculo $p$, y $q$ y $r$ son variables del cálculo, también es una aseveración la fórmula verifuncional obtenida al reemplazar $p$ por $q$, o por $\neg q$, o por $(q \vee r)$ en todas las posiciones que $p$ ocupa en $t$.

III Si $P$ y $Q$ son fórmulas verifuncionales tales que $\vdash P$ y $\vdash(\neg P \vee Q)$, entonces $\vdash Q$.

IV $\vdash(\neg(p \vee p) \vee p)$, $\vdash((\neg(p \vee (q \vee r)) \vee (q \vee (p \vee r)))$,
$\vdash(\neg q \vee (p \vee q))$, $\vdash(\neg(\neg q \vee r) \vee (\neg(p \vee q) \vee (p \vee r)))$,
$\vdash(\neg(p \vee q) \vee (q \vee p))$.[17]

15 Por ejemplo, cualesquiera que sean las fórmulas representadas por las letras $p$ y $q$, la expresión '$(p \vee q)$' representa la fila de signos formada por el paréntesis izquierdo, seguido de la fórmula representada por $p$, seguido del signo de disyunción, seguido de la fórmula representada por $q$, seguido del paréntesis derecho.

16 Post escribe "$p \vee q$", sin los paréntesis, que introduce luego sin estipulación previa en los esquemas de fórmulas más complejas. Pero el uso de los paréntesis tiene que regirse por un postulado, igual que el uso de los conectivos, aunque más tarde se autorice su eliminación —informal— si no hay riesgo de confusión.

17 Salvo por la sustitución de símbolos arriba descrita, reproduzco fielmente el Postulado IV de Post (p. 267). El lector poco familiarizado con la lógica moderna lo entenderá mejor si reemplaza cada fila de la forma $(\neg\alpha \vee \beta)$ por una de la forma $(\alpha \rightarrow \beta)$.

El *teorema fundamental* demostrado por Post acerca del cálculo proposicional establece la condición necesaria y suficiente para que una fórmula del conjunto definido mediante el postulado I "sea aseverada (*asserted*) en virtud de los postulados II, III y IV" (1921, p. 269), esto es —en la terminología del Apéndice IX— para que sea deducible de los axiomas (IV) conforme a las reglas de inferencia por sustitución (II) y *modus ponens* (III). La demostración de este teorema descansa en una interpretación algebraica de las fórmulas del cálculo (o de los esquemas con que Post los representa—no sabría decir cuál de estas dos alternativas es la que tiene presente; para no distraer al lector me referiré expresamente sólo a la primera). Dicha interpretación consiste sencillamente en esto: cada fórmula representa una operación algebraica sobre un dominio que contiene dos objetos. Post los designa con los signos '+' y '–', pero yo usaré los signos '0' y '1'. Como es habitual en este género de representaciones, las letras representan los argumentos de la operación. Así, una fórmula con *n* variables proposicionales diferentes representará una aplicación de $\{0,1\}^n$ en $\{0,1\}$. Como cabe esperar, la aplicación $\neg:\{0,1\} \rightarrow \{0,1\}$ se define por la doble condición $\neg 0 = 1$, $\neg 1 = 0$, y la aplicación $\vee:\{0,1\}^2 \rightarrow \{0,1\}$ por la cuádruple condición $(0 \vee 0) = (0 \vee 1) = (1 \vee 0) = 0$ y $(1 \vee 1) = 1$. Cada fórmula del cálculo proposicional (o cada esquema) representa entonces alguna aplicación $\varphi$: $\{0,1\}^n \rightarrow \{0,1\}$ definida por composición (generalmente repetida) de éstas dos. Post demuestra el teorema siguiente: Toda aplicación $\varphi$: $\{0,1\}^n \rightarrow \{0,1\}$ está representada por una fórmula.[18] Diremos con Post que una tal aplicación es *positiva* si su valor es idénticamente 0 en todos sus argumentos (recuérdese que Post escribe '+' en vez de '0'), *negativa* si su valor es idénticamente 1 ('–'), y *mixta* si admite ambos valores. Esta partición de las operaciones algebraicas sobre $\{0,1\}$ se extiende naturalmente a las fórmulas que las representan. Digamos con

18 La demostración es muy fácil de seguir si escribimos $(p \wedge q)$ —o, como Post, $p{\cdot}q$— para abreviar $\neg(\neg p \vee \neg q)$. En tal caso, la aplicación $\wedge$: $\{0,1\}^2 \rightarrow \{0,1\}$ queda definida por las ecuaciones $(0 \wedge 0) = 0$, $(0 \wedge 1) = (1 \wedge 0) = (1 \wedge 1) = 1$. La demostración procede por inducción sobre el número de argumentos *n*. Hay cuatro aplicaciones tales que $n = 1$, representadas, respectivamente, por $(p \vee p)$, $(p \vee \neg p)$, $\neg(p \vee \neg p)$ y $\neg p$. Supongamos que el teorema vale si $n = k$ y que $f$ aplica $\{0,1\}^{k+1}$ en $\{0,1\}$. Es fácil ver que $f(p_1, \dots, p_{k+1}) = ((p_{k+1} \wedge f_1(p_1, \dots, p_k)) \vee (\neg p_{k+1} \wedge f_2(p_1, \dots, p_k))$, donde $f_1$ y $f_2$ son aplicaciones de $\{0,1\}^k$ en $\{0,1\}$ y por ende están representadas por fórmulas, conforme a la hipótesis inductiva.

Post que una fórmula α *está contenida en* un fórmula β si α se deriva de β por una o más aplicaciones de la regla de sustitución II. Post demuestra que toda fórmula contenida en una fórmula positiva (negativa) es positiva (negativa) y que toda fórmula mixta con *n* variables diferentes contiene al menos una fórmula representativa de *cada* operación *n*-aria. *La condición necesaria y suficiente para que una fórmula verifuncional sea aseverada en virtud de los postulados II, III y IV es que ella sea positiva*. Es muy fácil demostrar que esta condición es necesaria: un cálculo rápido muestra que cualquier fórmula construida según los esquemas del postulado IV es positiva; ya señalamos que toda fórmula derivada por sustitución de una fórmula positiva es positiva, y es evidente que toda fórmula derivada por *modus ponens* de dos fórmulas positivas es positiva. Para demostrar que la condición es suficiente, Post establece un método general para derivar la aseveración ⊢α de cualquier fórmula positiva α. En otras palabras, Post nos da lo que se llama una demostración *constructiva*: no se limita a probar en abstracto que cada fórmula positiva es deducible por II y III de los axiomas IV, sino que enseña "*un método para escribir inmediatamente una derivación formal de su aseveración basada en los postulados*" (1921, p. 271; en cursiva en el original). La demostración de Post, relativamente simple, será el paradigma de otras más complejas. La reproduzco en el Apéndice XV.

Como la positividad o no positividad de una fórmula puede establecerse computando el valor de la operación respectiva para cada asignación de valores a sus argumentos, el teorema fundamental de Post resuelve el problema de la *d*-decisión para el cálculo proposicional. Pero su significación no se limita a eso. Como el lector se habrá dado cuenta, las fórmulas positivas según la interpretación algebraica son precisamente las fórmulas válidas según la interpretación lógica familiar presentada en el Apéndice IX.[19] Por lo tanto, al establecer que una fórmula verifuncional es deducible de los postulados II, III y IV si y sólo si es positiva, Post ha demostrado que el sistema deductivo descrito es a la vez *correcto* y *completo*, esto es, que toda fórmula

[19] Post mismo nos lo hace presente cuando introduce su interpretación algebraica del cálculo proposicional con estas palabras: "Denotaremos el valor veritativo de cualquier proposición *p* con + si es verdadera y con – si es falsa. Conviene tener en vista este significado de + y – como una guía del pensamiento, pero en la exposición siguiente deben considerarse meramente como símbolos que manipulamos de cierta manera" (1921, p. 267).

deducible es válida y toda fórmula válida es deducible (Apéndice IX.F).[20] Por esta razón, cualquier procedimiento de *v*-decisión —como el conocido método de las tablas de verdad utilizado en efecto por Post— constituye a la vez un procedimiento de *d*-decisión. Post establece además que el cálculo proposicional tiene la siguiente interesante propiedad: cualquiera que sea la formula $\alpha$, o bien $\alpha$ es deducible, o bien $\neg\alpha$ es deducible. Un sistema deductivo que posea esta propiedad —puramente sintáctica— suele llamarse *completo en el sentido de Post*.[21] Si $\Delta$ es el conjunto de todas las fórmulas deducibles en un sistema completo en este sentido y $\alpha$ es cualquier fórmula que no pertenece a $\Delta$, entonces $\Delta \cup \{\alpha\}$ es inconsistente (puesto que $\neg\alpha \in \Delta$). Por lo tanto, un sistema deductivo completo en el sentido de Post no puede fortalecerse sin engendrar contradicciones.[22]

20 Es oportuno señalar que Bernays, en su *Habilitationsschrift* inédito de 1918, había establecido que un sistema deductivo equivalente al considerado por Post es completo en el sentido indicado (toda tautología verifuncional es deducible de sus axiomas). Bernays 1926 es una versión abreviada de ese trabajo.

21 Conviene observar que el adjetivo 'completo' (*complete*) no se usa así en Post 1921. Allí, un "sistema completo" es un cálculo proposicional en que "every truth system has a representation", esto es, uno cuyos conectivos primitivos bastan para representar todas las aplicaciones de $\{V,F\}^n$ en $\{V,F\}$ para cualquier $n \geq 1$ (p. 273; véase arriba, la nota 18). En cambio, un sistema consistente, completo en el sentido de Post, en Post 1921 se llama "closed", *cerrado* (p. 277).

22 La mayor parte de la tesis doctoral de Post (1921) se consagra a generalizaciones del cálculo proposicional. Aunque caen fuera del tema de esta sección, diré aquí brevemente de qué se trata. Post considera generalizaciones de tres tipos: por variación de los conectivos primitivos, por variación de los postulados y por variación del número de objetos ("valores veritativos") contenidos en el dominio de la interpretación algebraica. Con respecto al primer tipo, conviene identificar los sistemas cuyos conectivos primitivos permiten representar el mismo conjunto de operaciones algebraicas sobre $\{0,1\}$; Post remite a un trabajo futuro en el cual promete demostrar que hay 66 sistemas diferentes cuyos conectivos primitivos son unarios, binarios y ternarios, y que, si admitimos conectivos primitivos *n*-arios con $n \geq 4$, se generan ocho familias infinitas de sistemas (cf. Post 1941); luego establece las condiciones que debe reunir los postulados de un sistema capaz de representar *todas* las operaciones sobre $\{0,1\}$ para que dicho sistema sea completo en el sentido de Post. A propósito del segundo tipo de generalización, Post estipula que llamará "inconsistente" a cualquier sistema en que sea deducible la variable proposicional *p*. El tercer tipo de generalización concierne a las llamadas "lógicas polivalentes"; Post ve una analogía entre el dominio de la interpretación algebraica de un cálculo proposicional y el espacio de una geometría, específicamente, entre la numerosidad de aquél y el número de dimensiones de éste; ella motiva la aguda observación siguiente:

En los próximos años Post trabajó intensamente en una teoría general de los sistemas formales (cálculos lógicos) con miras a resolver el problema de la *d*-decisión para el sistema entero de *Principia Mathematica*. Hacia 1924 se convenció —mediante un argumento "diagonal" como el utilizado por Turing (1937)— de que la solución tenía que ser negativa. Pero no publicó nada.[23] Por ese mismo tiempo, jóvenes matemáticos europeos, privados de la luz que podría haberles proporcionado un contacto siquiera verbal con su colega americano, acometían el *Entscheidungsproblem* con un entusiasmo digno de mejor causa. Concentraré mi atención en la obra de Herbrand, que constituye en cierto modo el *ne plus ultra* en esta materia. Su máximo logro es el Teorema de Herbrand a que he aludido un par de veces. Bernays (1954) lo llama "el teorema central de la lógica de los predicados". La tesis doctoral de Herbrand, *Investigaciones sobre la teoría de la demostración* (1930), está entera dirigida a probarlo. No es posible reproducir la prueba aquí: larguísima ya en la versión defectuosa del autor, ha crecido todavía más con las correcciones de Dreben y sus colaboradores.[24] Pero vale la pena explicar

---

> Mientras el espacio intuitivo de puntos con mayor número de dimensiones tiene tres, el espacio intuitivo de proposiciones con mayor número de dimensiones tiene dos. Pero tal como podemos interpretar intuitivamente los espacios geométricos de mayor número de dimensiones usando un elemento distinto del punto, interpretaremos los espacios de mayor número de dimensiones de nuestra lógica tomando como elemento algo que no sea la proposición.
>
> (Post 1921, p. 281)

23 Mucho más tarde, Post redactó un largo informe sobre sus investigaciones de ese período, al que incorpora extractos de sus notas, para documentar cómo había anticipado los grandes hallazgos de Gödel, Church y Turing a que me refiero en los Capítulos 2.10 y 2.11. Dicho informe fue publicado póstumamente en la antología *The Undecidable* editada por Martin Davis (1965).

24 Dos de las proposiciones auxiliares que Herbrand utiliza en su prueba y alegadamente demuestra son, en efecto, falsas: el decisivo lema del § 3.3 y el Lema 3 del § 5.3 (Herbrand 1930, cap. 5). Dreben, Andrews y Aanderaa 1963 dan contraejemplos que ponen su falsedad en evidencia. Dreben 1963 enuncia un nuevo lema con que reemplazar al del § 3.3. Una versión corregida del argumento de Herbrand se bosqueja en Dreben, Andrews y Aanderaa 1963a y Dreben y Anderaa 1964. Dreben y Denton 1966 dan una demostración detallada del lema sustituto. La admirable versión inglesa de los escritos de Herbrand editada por Goldfarb (Herbrand LW) contiene varias notas —en parte redactadas por Dreben— que explican claramente este asunto. Las notas de Dreben se publicaron ya con la traducción inglesa del capítulo 5 de Herbrand 1930 en van Heijenoort 1967.

el significado del Teorema.

Como vimos en la Sección 2.6.3, Herbrand considera teorías axiomáticas expresables en un cálculo predicativo de primer orden esencialmente igual a nuestro **CP1=** (las diferencias son muy superficiales y aquí las pasaré por alto). El enfoque es estrictamente sintáctico: aunque Herbrand habla de *propositions*, las caracteriza como filas de signos, por lo cual me parece justo llamarlas 'fórmulas'.[25] El cálculo está organizado como un sistema deductivo equivalente al sistema de primer orden de *Principia Mathematica*. Una fórmula *verdadera* (*vraie*) o *identidad* es una fórmula deducible de los axiomas conforme a las reglas de inferencia.

En una fase preliminar, Herbrand considera solamente fórmulas compuestas de letras minúsculas (*propositions-éléments*, correspondientes a nuestros predicados 0-ádicos y a las "variables proposicionales" de Whitehead y Russell), los conectivos $\neg$ y $\vee$, y signos de puntuación, esto es, lo que en el Apéndice IX.C llamo fórmulas verifuncionales. Con ellas organiza un sistema deductivo equivalente al definido por los Postulados I–IV de Post. Las fórmulas deducibles en este sistema son "identidades de la primera especie". Herbrand asocia a cada letra minúscula lo que llama un "valor lógico", consistente en una de las dos letras mayúsculas *V* y *F*. El valor lógico de una fórmula compuesta exclusivamente de variables proposicionales, conectivos y signos de puntuación depende de los valores lógicos asignados a sus variables proposicionales, según reglas que ya nos son familiares. Herbrand también designa los valores lógicos con las palabras francesas *vrai* y *faux* ('verdadero' y 'falso'), lo que puede ocasionar confusión con el uso de *vraie* en la acepción de 'deducible en el sistema'. Herbrand demuestra, como Post, que una fórmula verifuncional es deducible en el cálculo proposicional de *Principia Mathematica* si y sólo si su valor lógico es *V* cualquiera que sea el valor lógico asignado a las minúsculas componentes.

Herbrand pasa luego a ocuparse con su tema principal, el cálculo predicativo de primer orden con identidad, que, para mayor claridad, identificaré con la versión máxima de nuestro **CP1=** (con $\neg$ y $\vee$ como únicos conectivos "oficiales"). Forma este cálculo como una extensión del anterior, agregándole los signos de cuantificación y de igualdad y secuencias infinitas de (i)

[25] "Una colección (*assemblage*) de letras, signos y puntos formada a partir de los signos $\sim$ y $\vee$ conforme a las reglas precedentes y a las que se verán en 1.4 se llama una *proposición*" (Herbrand 1930, cap. 1, § 1.3; EL, p. 44; cf. p. 72).

variables individuales, (ii) constantes individuales, (iii) predicados *n*-ádicos, para cada $n \geq 1$ y (iv) functores *n*-arios, para cada $n \geq 1$. Herbrand trata a las fórmulas sin variables individuales que este cálculo hereda del anterior como "predicados con 0 términos" (*fonctions propositionelles à 0 argument*— EL, pp. 59f.; el lector reconocerá aquí la fuente de nuestra expresión 'predicado 0-ádico'). Una *proposition élément* consta de un predicado seguido del número apropiado de variables. Me permitiré llamarla *fórmula elemental*, como en el Apéndice IX. Por su parte, Herbrand llama *propositions élémentaires* a las fórmulas compuestas de *propositions éléments* y que no contienen cuantificadores; pero a estas fórmulas —que Herbrand también describe como *fonctions propositionelles de première espèce*— yo las llamaré *matrices*. Los *componentes elementales* de una matriz son las distintas fórmulas elementales que figuran en ella. Llamaré *valuación* de una matriz a la asignación arbitraria de un "valor lógico" *V* o *F* a cada uno de sus componentes elementales; cada valuación confiere obviamente un "valor lógico" determinado a la matriz misma, computable por el procedimiento de las tablas de verdad.

Si admitimos como axiomas a todas las fórmulas construidas sustituyendo letras por fórmulas elementales en los esquemas del Postulado IV de Post, el teorema relativo a las identidades de la primera especie se extiende naturalmente a las matrices: una matriz se deduce de dichos axiomas por sustitución y *modus ponens* si y sólo si su valor es *V* en todas sus valuaciones. En tal caso, decimos con Herbrand que la matriz es una identidad de la primera especie con respecto a sus componentes elementales (obviamente, una matriz $\mu$ tiene esta propiedad si y sólo si $\mu$ es una tautología en el sentido del Apéndice IX.C). Como esta propiedad es decidible mediante un algoritmo, Herbrand se desentiende de los axiomas tradicionales del cálculo proposicional y adopta como axiomas del cálculo de primer orden *todas las matrices* cuyo valor sea *V* en cualquier valuación.[26] No admite otros axio-

[26] "*Les propositions élémentaires qui, considérées comme des fonctions propositionelles de première espèce de leurs propositions-éléments, sont des identités de première espèce seront vraies*" (Herbrand 1930, cap. 2, § 2.2,; EL, p. 63; cursiva en el original). Se advertirá que esta clase de fórmulas es más estrecha que la de las tautologías, definida en el Apéndice IX.C, pues comprende sólo aquéllas de entre éstas que no contengan cuantificadores. Por ejemplo, la tautología $(\forall xPx \vee \neg\forall xPx)$ no es un axioma según la estipulación citada.

mas. Adopta las siguientes reglas de inferencia: 1º *modus ponens*; 2º la regla de simplificación: de $(\phi \vee \phi)$, deducir $\phi$; 3º las reglas de generalización—si $\phi(x)$ contiene la variable libre $x$, de $\phi(x)$ deducir $\forall x\phi(x)$ o $\exists y\phi(x,y)$, donde $\phi(x,y)$ es cualquier fórmula obtenida reemplazando $x$ por $y$ en *una o más* posiciones libres de $x$ en $\phi(x)$; 4º las reglas de transición (*règles de passage*) que autorizan a sacar un cuantificador del alcance de un conectivo o a ponerlo dentro de él (conforme a los esquemas (4), (5), (8) y (9) del Apéndice XIII). Herbrand da por supuesto que las variables ligadas son intercambiables;[27] específicamente, una variable $\xi$ se puede reemplazar por otra variable cualquiera $\zeta$ en todas las posiciones en que $\xi$ esté ligada por un cierto cuantificador en una fórmula $\phi$, siempre que $\zeta$ no figure libre dentro del alcance de ese cuantificador ni esté ligada en $\phi$ por otro cuantificador cuyo alcance incluya posiciones de la variable que se quiere reemplazar. Pero Herbrand no postula expresamente una regla de inferencia a este efecto.

Herbrand dice que "el problema fundamental de la lógica matemática consiste en hallar un procedimiento que permita siempre reconocer si una fórmula (*proposition*) es deducible (*vraie*) o no en esta teoría" (1930, cap. 2, 2.1; EL, p. 64). Su estrategia para resolverlo consistió en tratar de reducir el caso general al caso especial de las matrices, ya resuelto. Para ello asocia una clase bien determinada de matrices a cada fórmula del cálculo. Explicaré cómo construir tales matrices en el caso de una fórmula cualquiera $\phi$ de nuestro **CP1=**. Divido la explicación en varios pasos, numerados para mayor claridad.

[27] "Se sobreentiende que la letra que designa una variable puede ser reemplazada sin inconveniente por otra (y, en particular, se puede utilizar cualquier letra como variable ligada), a condición eso sí de que dos variables [dentro de una misma fórmula] no pueden designarse con la misma letra más que en el caso en que ambas estén ligadas y que sus alcances [en nuestra jerga: el alcance de los cuantificadores que las ligan—R.T.] no tengan ningún signo en común" (Herbrand, EL, p. 62). A la luz de la equivalencia (7) del Apéndice XIII es claro que esta convención está semánticamente justificada en nuestro **CP1=**. Sin ella no es posible deducir la tautología $(\forall xPx \vee \neg\forall xPx)$ en el sistema de Herbrand. En efecto, partiendo del axioma $(Px \vee \neg Px)$ tenemos la deducción siguiente (donde simbolizo con $\Rightarrow$ cada aplicación de una de las reglas de generalización o transición): $(Px \vee \neg Px) \Rightarrow \exists y(Px \vee \neg Py) \Rightarrow (Px \vee \exists y\neg Py) \Rightarrow (Px \vee \neg\forall yPy) \Rightarrow \forall x(Px \vee \neg\forall yPy) \Rightarrow (\forall xPx \vee \neg\forall yPy)$. Conforme a la convención citada, la conclusión puede reescribirse $(\forall xPx \vee \neg\forall xPx)$.

(1) Como las variables ligadas son intercambiables en la forma arriba especificada, podemos, sin pérdida de generalidad, requerir que $\phi$ cumpla las tres condiciones siguientes: ninguna variable libre en $\phi$ ocupa en $\phi$ posiciones donde está ligada; toda variable ligada en $\phi$ figura en el alcance del cuantificador que la liga; no hay en $\phi$ dos cuantificadores que liguen la misma variable. En el Apéndice XIII llamamos *regular* a una fórmula que cumple estas condiciones. Supondremos, pues, que $\phi$ es una fórmula regular.

(2) Llamamos *prenexa* a una fórmula en la cual ningún cuantificador figura dentro del alcance de un conectivo (Apéndice XIII). Llamaré *antiprenexa* a una fórmula en la cual ningún conectivo figura dentro del alcance de un cuantificador. Aplicando las reglas de transición en uno o el otro sentido, se deduce de $\phi$ una fórmula prenexa $\phi^*$ y una fórmula antiprenexa $\phi^\dagger$. Como es obvio, $\phi$ puede deducirse a su vez de cualquiera de estas dos fórmulas. Diré, por ello, que cualquiera de estas tres fórmulas *equivale deductivamente* a las otras dos. Herbrand llama a $\phi^\dagger$ la *forma canónica* de $\phi$.

(3) Asignamos a cada fórmula $\psi$ una *altura*, determinada por los functores que contiene, de acuerdo con las reglas siguientes: (i) Si $\psi$ no contiene functores su altura es 0; de otro modo, la altura de $\psi$ es igual a la máxima altura que tenga un functor en $\psi$. (ii) La altura de un functor $\mathfrak{f}$ en una determinada posición dentro de la fórmula $\psi$ depende de su alcance en esa posición: si éste no contiene functores, $\mathfrak{f}$ tiene altura 1; pero si el alcance de $\mathfrak{f}$ contiene un functor de altura $k$ y no contiene ningún functor de altura mayor que $k$, la altura de $\mathfrak{f}$ en la posición considerada es $k+1$. Sea, pues, $h \geq 0$ la altura de nuestra fórmula $\phi$.

(4) Sea $\xi$ una variable ligada en $\phi$ por cierto cuantificador Q. Diremos con Herbrand que $\xi$ es una variable *general* si Q es universal y está dentro del alcance de un número par de signos de negación o si Q es existencial y está dentro del alcance de un número impar de signos de negación, y que $\xi$ una variable *restringida* si Q es existencial y está dentro del alcance de un número par de signos de negación o si Q es universal y está dentro del alcance de un número impar de signos de negación. El lector podrá comprobar que $\xi$ es en $\phi$, respectivamente, una variable general o restringida, según que el cuantificador que liga a $\xi$ en la fórmula prenexa $\phi^*$ sea universal o existencial.

(5) Supongamos que $\xi$ y $\zeta$ son dos variables distintas ligadas respectivamente en $\phi$ por los cuantificadores Q y Q′. Si el alcance de Q′ está contenido en el

alcance de Q diremos que $\xi$ *supera* a $\zeta$ en $\phi$. Si $\xi$ supera a $\zeta$ en la fórmula antiprenexa $\phi^\dagger$ (la forma canónica de $\phi$) diremos que $\xi$ *domina* a $\zeta$ en $\phi$.

(6) Sea $\zeta$ una variable general dominada en $\phi$ por $r$ variables restringidas $\xi_1, \ldots, \xi_r$ ($r > 0$). Asociamos a $\zeta$ de modo exclusivo un functor $r$-ario $\mathfrak{f}_\zeta$, elegido entre los functores del **CP1=** que no figuran en $\phi$. La exclusividad requerida está garantizada si numeramos las variables ligadas de $\phi$ según el orden en que aparecen —de izquierda a derecha— en esa fórmula y los functores $r$-arios que no figuran en $\phi$ según el número de palotes que contienen, y asociamos el $k$-ésimo de estos functores $r$-arios a la $k$-ésima variable general dominada en $\phi$ por $r$ variables restringidas. Si la variable general $\zeta$ no está dominada por ninguna variable restringida le asociamos, también en forma exclusiva, una constante que no figure en $\phi$.

(7) Ahora explicaré, siguiendo de cerca a Herbrand, cómo se construye, para cada entero positivo $p$, un conjunto finito de objetos $\Delta(\phi,p)$ dependiente de nuestra fórmula $\phi$. Los elementos de $\Delta(\phi,p)$ serán constantes del **CP1=** que figuran en $\phi$ o han sido asociadas a una variable general de $\phi$ no dominada por variables restringidas, o provendrán de la secuencia $\mathfrak{K} = \kappa_1, \kappa_2, \ldots$ de las constantes del **CP1=** (ordenadas por el número de palotes) que no están comprendidas entre las anteriores. Tendremos que referirnos a más de una secuencia formada eliminando de $\mathfrak{K}$ todas las constantes que pertenecen a algún conjunto $C$ y renumerando las restantes en el orden en que figuran en $\mathfrak{K}$; llamaré $\mathfrak{K}\backslash C$ a la secuencia así definida. Primero construiremos una lista $C_1, \ldots, C_{p+1}$ de conjuntos de constantes que figuran en $\mathfrak{K}$, como sigue: $C_1$ es el conjunto de las constantes que figuran en $\phi$ o han sido asociadas a una variable general de $\phi$. Si $\phi$ no contiene constantes ni variables generales no dominadas por variables restringidas, estipulamos que $C_1 = \{\kappa_1\}$.[28] Para construir $C_2$ procedemos así: Con los functores de altura 1 que figuran en $\phi$ y los functores asociados a las variables generales de $\phi$ formamos una lista de functores $\mathfrak{F} = \mathfrak{f}_1, \ldots, \mathfrak{f}_s$ ordenada lexicográficamente por la $r$-aridad y el número de palotes de cada uno; si $\mathfrak{f}_t$ es $m$-ario ($1 \leq t \leq s$; $m > 0$), le asignamos, en forma exclusiva, una constante de la secuencia $\mathfrak{K}$ al par formado por $\mathfrak{f}_t$ y cada $m$-tuplo $\mathbf{k} \in C_1{}^m$. Concebimos esta constante como el "valor" de $\mathfrak{f}_t$ en dicho $m$-tuplo. Tales valores se eligen de modo que constituyan

28 De hecho, en ese caso $\mathfrak{K}$ comprende *todas* las constantes del **CP1=**, $a$, $a_|$, $a_{||}$, . . . ; así que la estipulación antedicha significa que $C_1 = \{a\}$.

precisamente un segmento inicial $\kappa_1, \ldots, \kappa_n$ de la secuencia $\mathfrak{K} \setminus C_1$. Entonces, $C_2 = \{\kappa_1, \ldots, \kappa_n\}$, a menos que $\kappa_1 \in C_1$, en cuyo caso $C_2 = \{\kappa_2, \ldots, \kappa_n\}$. Supongamos ahora que se han construido las colecciones $C_1, \ldots, C_k$ ($k < p$). Sea $U_k = C_1 \cup \ldots \cup C_k$ . Entonces $C_{k+1}$ estará formado por constantes de la secuencia $\mathfrak{K}$ elegidas entre las que no pertenecen a $U_k$ de tal modo que (i) si $\kappa \in C_{k+1}$, $\kappa$ es el valor de un determinado functor $m$-ario $\mathfrak{f}_t$ de la lista $\mathfrak{F}$ en un determinado $m$-tuplo $\lambda \in U_k{}^m$, tal que $\lambda \notin U_{k-1}{}^m$;[29] (ii) los elementos de $U_{k+1}$ que no pertenecen a $C_1$ constituyen un segmento inicial de $\mathfrak{K}$. Estipulamos que $U_p = C_1 \cup \ldots \cup C_p = \Delta(\phi,p)$. Sea $N$ la cardinalidad del conjunto $\Delta(\phi,p)$. Los elementos de $\Delta(\phi,p)$ forman una subsecuencia finita $\alpha_1, \ldots, \alpha_N$, de la secuencia $a$, $a_{|}$, $a_{||} \ldots$, de las constantes del **CP1=**.

(8) Asociaré a $\phi$ una matriz $\phi^R$ que llamo, con Herbrand, la reducida (*réduite*) de $\phi$.[30] La defino por inducción sobre las subfórmulas de $\phi$. Si $\psi$ es una fórmula elemental, su reducida $\psi^R = \psi$. Si $\psi$ y $\chi$ son fórmulas cualesquiera, $(\neg\psi)^R = \neg(\psi^R)$ y $(\psi \vee \chi)^R = (\psi^R \vee \chi^R)$. Si $\xi$ es una variable general dominada en $\phi$ por las variables restringidas $\zeta_1, \ldots, \zeta_k$ y la subfórmula $\psi$ es el alcance del cuantificador Q$\xi$ que liga a $\xi$ en $\phi$, $(\mathrm{Q}\xi\psi)^R$ es la fórmula que se obtiene reemplazando $\xi$ por el término $\mathfrak{f}_\xi\zeta_1 \ldots \zeta_k$ en todas las posiciones libres de $\xi$ en $\psi^R$ (recuérdese que $\mathfrak{f}_\xi$ es el functor $k$-ario asociado a la variable general $\xi$; como las variables $\zeta_1, \ldots, \zeta_k$ dominan a $\xi$, están libres en todas las posiciones que ocupan en Q$\xi\psi$). Si $\xi$ es una variable restringida y $\psi$ es el alcance del cuantificador que liga a $\xi$ en $\phi$, designo con $\psi^R(\alpha/\xi)$ a la fórmula resultante de reemplazar la variable $\xi$ por la constante $\alpha$ en todas las posiciones de $\xi$ en $\psi^R$ (en virtud de nuestras convenciones iniciales, todas esas posiciones son libres). Como sabemos, el cuantificador que liga a $\xi$ será existencial (si está en el alcance de un número par de negaciones) o universal (si está en el alcance de un número impar de negaciones). Según cual sea el caso, aplicamos una de las dos convenciones siguientes:[31]

29 Esto quiere decir que si $\lambda = \langle\lambda_1, \ldots, \lambda_m\rangle$, hay al menos un entero positivo $j \leq m$ tal que $\lambda_j \in C_k$. La estipulación sirve para asegurar que $\mathfrak{f}_t$ no tenga en $C_{k+1}$ nuevos valores correspondientes a $m$-tuplos formados con elementos de $U_{k-1}$, a los que ya se asignó un valor de $\mathfrak{f}_t$ en $U_k$.

30 En la citada versión inglesa —Herbrand LW— *réduite* se traduce *expansion*.

31 En la segunda igualdad uso como Herbrand, en aras de la claridad, el conectivo $\wedge$; pero se sobreentiende que $\psi \wedge \chi$ no es más que una abreviatura de $\neg(\neg\psi \vee \neg\chi)$.

$$(\exists\xi\psi)^{R} = \psi^{R}(\alpha_1/\xi) \vee \ \ldots \ \vee \psi^{R}(\alpha_N/\xi)$$

$$(\forall\xi\psi)^{R} = \psi^{R}(\alpha_1/\xi) \wedge \ \ldots \ \wedge \psi^{R}(\alpha_N/\xi)$$

(donde $\{\alpha_1, \ldots ,\alpha_N\} = \Delta(\phi,p)$). La construcción descrita determina la reducida $\phi^R$ de nuestra fórmula $\phi$ si ésta no contiene variables libres. Si $\phi$ contiene las variables libres $\eta_1, \ldots , \eta_r$, $\phi^R$ será, por definición, la reducida de su clausura universal: $\phi^R = (\forall\eta_1 \ldots\forall\eta_r\phi)^R$. Recordando que la altura de $\phi$ es $h$, vemos que la matriz $\phi^R$ contiene functores de altura $q \leq h$.

(9) Ahora asociaremos a $\phi$ una nueva matriz que no contenga functores. Observamos que el alcance de cada functor $m$-ario $\mathfrak{f}$ que ocupa en $\phi^R$ una posición de altura 1 consiste precisamente en un $m$-tuplo de constantes pertenecientes a $\Delta(\phi,p) = U_p$; en virtud de la construcción del párrafo anterior, $U_{p+1}$ contiene una constante $\alpha$ que es el valor de $\mathfrak{f}$ en dicho $m$-tuplo. Reemplazamos por $\alpha$ el término que $\mathfrak{f}$ encabeza en la posición considerada. Para reemplazar en forma análoga los functores de altura superior a 1, construimos los conjuntos de constantes $C_{p+2}, \ldots , C_{p+h}$ definidos como sigue: si $\mathfrak{f}$ es un functor $m$-ario que tiene altura $q$ en cierta posición que ocupa en $\phi$ $(1 < q \leq h)$, $C_{p+q}$ contiene un valor para cada argumento admisible de $\phi$ en $(U_{p+q-1})^m$; dicho valor pertenece a $\mathfrak{K}\setminus U_{p+q-1}$ y precede a todos los elementos de $\mathfrak{K}\setminus U_{p+q}$. Los "argumentos admisibles" en cuestión son todos los $m$-tuplos $\langle\lambda_1, \ldots ,\lambda_m\rangle \in (U_{p+q-1})^m$ determinados así: si el $j$-ésimo término en el alcance de $\mathfrak{f}$ en la posición considerada es una constante, $\lambda_j$ es esa constante; si es una variable, $\lambda_j$ puede ser cualquiera de las constantes de $\mathfrak{K}$ incluidas en $U_{p+1}$; si comienza con un functor de altura $w < q$, $\lambda_j$ puede ser cualquiera de los valores asignados a ese functor en $U_{p+w}$. Los functores de altura mayor que 1 pueden ahora eliminarse en $h - 1$ etapas sucesivas. Al comienzo de la $q$-ésima etapa $(1 < q \leq h)$, cada functor $m$-ario $\mathfrak{f}$ que ocupa en $\phi^R$ una posición de altura $q$ va seguido en esa posición por un $m$-tuplo de constantes pertenecientes a $U_{p+q-1}$; el término encabezado por $\mathfrak{f}$ se reemplaza entonces por el valor de $\mathfrak{f}$ en dicho $m$-tuplo, el cual es una constante perteneciente a $U_{p+q}$. El resultado final es una matriz sin functores que llamaré la matriz de orden $p$ para controlar la deducibilidad de $\phi$, o, abreviadamente, la $(d,p)$-matriz de $\phi$. La designo $\Pi_d(\phi,p)$.

(10) Herbrand enseña también a construir lo que llamaré la matriz de orden $p$ para controlar la consistencia de $\phi$, o $(c,p)$-matriz de $\phi$. La designo $\Pi_c(\phi,p)$.

Los pasos (6)–(9) de la construcción de $\Pi_d(\phi,p)$ describen exactamente la construcción de $\Pi_c(\phi,p)$ si intercambiamos los términos 'variable general' y 'variable restringida' (de tal modo que ahora se asocie un functor *r*-ario $\mathfrak{f}_\zeta$ a cada variable *restringida* dominada por *r* variables *generales*, etc.).[32] La colección de constantes construida por esta vía en el paso (7) normalmente difiere de $\Delta(\phi,p)$; si hay que referirse a ella podemos llamarla $\Gamma(\phi,p)$.

Por fin estamos en condiciones de enunciar el *Teorema de Herbrand*: la fórmula $\phi$ es deducible (*vraie*) si y sólo si hay un entero positivo *p* tal que $\Pi_d(\phi,p)$ es una identidad de la primera especie con respecto a sus componentes elementales. La demostración del Teorema enseña, de paso, a deducir de los axiomas, sin invocar la regla *modus ponens*, cualquier fórmula $\phi$ tal que, para algún $p > 0$, $\Pi_d(\phi,p)$ cumpla la condición antedicha.[33] El Teorema se puede enunciar también, como es obvio, haciendo referencia a las matrices para controlar la consistencia: la fórmula $\phi$ es consistente (con los axiomas del sistema) si y sólo si para cada entero positivo *p* hay una valuación

[32] Esta correspondencia en apariencia mágica entre los procedimientos para construir $\Pi_d(\phi,p)$ y $\Pi_c(\phi,p)$ tiene una explicación muy sencilla. $\phi$ es consistente con los axiomas del sistema si y sólo si su negación $\neg\phi$ no se deduce de ellos. Así, la matriz para controlar la deducibilidad de $\neg\phi$ servirá también para controlar la consistencia de $\phi$: $\Pi_d(\neg\phi,p) = \neg\Pi_c(\phi,p)$. Ahora bien, la forma antiprenexa $(\neg\phi)^\dagger = \neg(\phi^\dagger)$, de suerte que las relaciones de dominio entre variables son las mismas en $\phi$ y $\neg\phi$; pero la forma prenexa $(\neg\phi)^*$ difiere $\neg(\phi^*)$ precisamente en cuanto las variables ligadas en $\neg(\phi^*)$ por un cuantificador universal están ligadas en $(\neg\phi)^*$ por un cuantificador existencial y viceversa (cf. Apéndice XIII, (4) y (5)), de suerte que las variables generales y restringidas de $\phi$ son, respectivamente, las variables restringidas y generales de $\neg\phi$.

[33] Por lo tanto, el famoso Teorema del Corte de Gentzen (1936), que enseña a deducir cualquier fórmula prenexa sin aducir el equivalente gentzeniano de *modus ponens*, es un caso especial del Teorema de Herbrand. La expresión '$(\Pi_d(\phi,p)$' es mía; Herbrand enuncia su teorema en términos de "la propiedad B de orden *p*" definida así: La fórmula $\phi$ tiene la propiedad B de orden *p* si (lo que yo he llamado) la $(d,p)$-matriz de $\phi$ es una identidad de la primera especie con respecto a sus componentes elementales. El Teorema dice entonces que "la condición necesaria y suficiente para que una proposición sea deducible (*vraie*) es que tenga la propiedad B [de orden *p*, para algún entero positivo *p*]" (Herbrand, EL, p. 139; el texto entre corchetes no corresponde a nada que figure en el original francés, pero es indispensable; la versión inglesa contiene una inserción equivalente). La elección de la letra B para nombrar la propiedad en cuestión se justifica porque en la demostración del Teorema Herbrand apela a otras dos propiedades de fórmulas, definidas por él, que llama A y C.

de la matriz $\Pi_c(\phi,p)$ tal que el valor lógico de ésta es $V$.[34]

El siguiente ejemplo ilustra la fuerza del Teorema. Sea $\phi$ una fórmula que no contiene functores y equivale deductivamente a una fórmula prenexa en la que ningún cuantificador existencial precede a un cuantificador universal. Entonces hay un algoritmo para determinar si $\phi$ es o no deducible. En efecto, en tal caso las variables generales de $\phi$ no están dominadas por ninguna variable restringida, de modo que, para todo entero positivo $p$, $C_{p+1} = \varnothing$ y $\Delta(\phi,p) = \Delta(\phi,1) = C_1$. Por lo tanto, $\phi$ es deducible si y sólo si $\Pi_d(\phi,1)$ es una identidad de la primera especie con respecto a sus componentes elementales. Basta, pues, aplicar a $\Pi_d(\phi,1)$ el algoritmo que determina si esta matriz es o no una tautología, para decidir si $\phi$ es o no deducible. El *Entscheidungsproblem* para esta clase de fórmulas había sido resuelto por Bernays y Schönfinkel (1928), pero la solución basada en el Teorema de Herbrand es enormemente más simple. Por otra parte, el argumento que acabo de dar sugiere que el *Entscheidungsproblem* bien podría ser insoluble en el caso general. En efecto, si $\phi$ es una fórmula cualquiera podría ocurrir que —en contraste con la clase de fórmulas recién considerada— $\Delta(\phi,p+1)$ contenga más elementos que $\Delta(\phi,p)$, para cada $p > 0$. Entonces, aunque se determine mediante un algoritmo que $\Pi_d(\phi,p)$ no es una tautología para cierto $p$, no cabe excluir que $\Pi_d(\phi,q)$ lo sea para algún $q > p$ y el procedimiento de decisión puede prolongarse indefinidamente.

Herbrand no se hace cargo expresamente de esta última posibilidad, pero usa una terminología que inescapablemente —me parece— lleva a pensar en ella. Dice, en efecto, que una fórmula $\phi$ es "verdadera en un campo infinito (*champ infini*)" si cada matriz de control de consistencia $\Pi_c(\phi,p)$ admite una valuación que le asigna el valor lógico $V$, y que es "falsa en un campo infinito" si cada matriz de control de validez $\Pi_d(\phi,p)$ admite una valuación que le asigna el valor lógico $F$ ($p > 0$). Según esto, $\phi$ es verdadera en un campo infinito si y sólo si $\neg\phi$ es falsa en un campo infinito y $\phi$ es falsa en un

34 Como recordé en la nota 32, la fórmula $\phi$ es consistente si y sólo si $\neg\phi$ no es deducible. Según el primer enunciado del Teorema, $\neg\phi$ es deducible si y sólo si hay un $p > 0$ tal que $\Pi_d(\neg\phi,p)$ es una tautología. $\Pi_d(\neg\phi,p) = \neg\Pi_c(\phi,p)$. Ahora bien, si no hay un $p > 0$ tal que $\neg\Pi_c(\phi,p)$ sea una tautología, tenemos que, para cada $p > 0$ hay una valuación de $\neg\Pi_c(\phi,p)$ y $\Pi_c(\phi,p)$ que asigna a $\neg\Pi_c(\phi,p)$ el valor $F$ y, por ende, el valor $V$ a $\Pi_c(\phi,p)$.

campo infinito si y sólo si $\neg\phi$ es verdadera en un campo infinito.[35] Con este vocabulario, nuestro segundo enunciado del Teorema de Herbrand puede expresarse con gran concisión: $\phi$ es consistente (con los axiomas del sistema) si y sólo si es verdadera en un campo infinito. Equivalentemente, $\phi$ es deducible si y sólo si no es falsa en un campo infinito. En palabras de Herbrand:

> TEOREMA 1. *Si P es una identidad, ~P no puede ser verdadera en un campo infinito.*
> TEOREMA 2. *Si P no es una identidad, se puede fabricar un campo infinito en el que ~P es verdadera.*
>
> (Herbrand, EL, p. 143)

Herbrand observa que "Löwenheim (1915) ya enunció resultados análogos", pero le hace dos reproches. En primer lugar, Löwenheim confiere a la noción de 'verdad en un campo infinito' un significado "intuitivo", por lo cual "su demostración del Teorema 2 no alcanza el rigor que nos parece deseable". Pero además, "y éste es el reproche más grave, Löwenheim considera, al parecer, justamente a causa del sentido intuitivo que le da a dicha noción, que el Teorema 1 es evidente". Según Herbrand, ello es "completamente inadmisible" y llevaría, por ejemplo, "a considerar evidente la no-contradicción de la aritmética" (EL, p. 143). Este reproche nace de un malentendido: Löwenheim usa la palabra 'identidad', como es habitual en matemáticas, para referirse a una ecuación que está satisfecha como quiera que se le asignen valores a sus variables. Las identidades de primer orden de su cálculo de relativos, traducidas a nuestro **CP1=**, son precisamente las fórmulas *válidas*, es decir, verdaderas en cualquier interpretación. Si 'identidad' se entiende de este modo, el citado Teorema 1 es trivial. Pero lo que Herbrand llama 'identidad' es una fórmula *deducible* en su cálculo; se trata, pues, de un término de significado puramente sintáctico, no semántico como el de Löwenheim. El equívoco oculta una relación importante; tomados simultáneamente en su doble acepción, sintáctica y semántica, los Teoremas 1 y 2 caracterizan a la vez las fórmulas deducibles y las fórmulas válidas: $\phi$ es

35 Puesto que $\Pi_c(\phi,p) = \neg\Pi_d(\neg\phi,p)$ y $\Pi_c(\neg\phi,p) = \neg\Pi_d(\neg\neg\phi,p)$, y, como es obvio, $\Pi_d(\neg\neg\phi,p)$ equivale —semántica y deductivamente— a $\Pi_d(\phi,p)$.

deducible en el sistema de Herbrand si y sólo si hay un entero $p > 0$ tal que $\Pi_c(\neg\phi,p)$ tenga el valor $F$ en cada valuación, y esto ocurre si y sólo si $\phi$ es válida. Por lo tanto, si, desestimando los reproches de Herbrand, considerásemos que en su versión semántica el Teorema 1 es obvio y el Teorema 2 fue demostrado por Löwenheim,[36] concluiríamos que al establecerlos en su versión sintáctica Herbrand ha probado que su sistema deductivo para el cálculo predicativo de primer orden es *correcto* y *completo* (Apéndice IX.F). Como veremos enseguida (en el Capítulo 2.8), Kurt Gödel se doctoró en matemáticas casi al mismo tiempo que Herbrand, con una tesis en la que demuestra —sin equívocos— que dicho sistema es completo, esto es, suficiente para deducir todas las fórmulas válidas de primer orden. Pero Gödel adopta de entrada el punto de vista semántico —con lucidez sin paralelo hasta entonces— y en el punto decisivo de la prueba avanza al infinito sin miedo, "con arreglo a formas de inferencia conocidas" (*nach bekannten Schlußweisen*—Gödel, CW, I, 116), propias de la matemática clásica. Gracias a ello, produce un escrito de gran claridad y limpieza, digno antecesor de la obra revolucionaria (Gödel 1931) que estudiaremos en el Capítulo 2.10.[37]

36 Para que el Teorema 2 arriba enunciado sea un corolario del Teorema 2 de Löwenheim 1915 no sólo hay que entender 'identidad' en su acepción semántica, sino además 'campo infinito' en su sentido literal y no, como Herbrand, sólo como una *façon de parler*. Como se vio en la p. 252, el Teorema 2 de Löwenheim 1915 dice que toda ecuación huidiza de primer orden falla para alguna asignación de valores a sus variables en un dominio enumerable. Una ecuación huidiza en el sentido de Löwenheim corresponde en nuestro **CP1=** a una fórmula $\phi$ tal que (i) $\phi$ no es válida, pero (ii) $\neg\phi$ es falsa en cualquier interpretación $\langle\mathcal{D}_p,f_p\rangle$ tal que $|\mathcal{D}_p|$ es un entero positivo $p$. El Teorema 2 de Löwenheim 1915 implica que hay una interpretación $\langle\mathcal{D}_\omega,f_\omega\rangle$ tal que $|\mathcal{D}_\omega| = \omega$ y $f_\omega(\phi) = 1$, de modo que $\phi$ es falsa y, por ende, $\neg\phi$ es verdadera en dicha interpretación. (Si $\phi$ no es válida pero corresponde a una ecuación huidiza, $\neg\phi$ no satisface la condición (ii) y por lo tanto es realizable en un dominio finito y, *a fortiori*, en cualquier dominio infinito que lo contenga). En su luminosa introducción a la traducción inglesa de las obras de Herbrand, Goldfarb explica cómo la construcción de lo que llamé matrices de control viene a ser "un sustituto finitista (*a finitistic surrogate*)" de la noción semántica de realizabilidad en un dominio enumerable (Herbrand, LW, p. 11; cf. pp. 11–14).

37 No debo cerrar este capítulo sin mencionar otro giro adoptado por Herbrand para comunicar su hallazgo: el *Entscheidungsproblem* se reduce en cada caso "a un problema que es una generalización simple del problema de la resolución efectiva de un sistema de ecuaciones diofánticas" (EL, p. 199; cf. pp. 31s., 193ss.) Una ecuación *diofántica*

es una ecuación algebraica cuyos coeficientes son números enteros y cuyas soluciones se buscan entre los enteros. La observación de Herbrand implica que una solución del *Entscheidungsproblem* resolvería de paso el 10º problema de Hilbert: Dada una ecuación con cualesquiera incógnitas y coeficientes enteros, indicar un procedimiento que permita decidir mediante un número finito de operaciones si esa ecuación admite números enteros como soluciones (Hilbert 1900a en GA, III, p. 310). Matijasevic (1970) se apoyará en la solución negativa del *Entscheidungsproblem* por Church y Turing (*vide infra*, Capítulo 2.11) para demostrar que el 10º problema de Hilbert es asimismo insoluble. Y. I. Manin, en su *Curso de lógica matemática* (1977), ofrece una excelente explicación del descubrimiento de Matijasevic (cap. VI), que precede a su exposición de los teoremas de incompletud de Gödel (cap. VII) y de la solución negativa del *Entscheidungsproblem* (cap. VIII).

## 2.8 EL CÁLCULO PREDICATIVO DE PRIMER ORDEN ES COMPLETO

En este capítulo estudiaremos el primero de los grandes descubrimientos de Gödel: el cálculo predicativo de primer orden, organizado como sistema deductivo a la manera de Frege, o de Whitehead y Russell, es un *cálculo completo*, en el sentido preciso que explicaré enseguida y que, siguiendo justamente a Gödel, le damos comúnmente a esta expresión (v. Apéndice IX.F). Gödel enuncia y demuestra este resultado en su tesis doctoral (1929) y también en un artículo publicado en la revista de Hilbert, *Mathematische Annalen* (1930). Casi más importante que la demostración —que ha sido reemplazada generalmente en la literatura didáctica por otra más versátil y expedita debida a Henkin (1949)— es la clara concepción de las ideas envueltas, especialmente el distingo entre *validez* (o verdad lógica) y *deducibilidad*. Con ello se consolida el enfoque semántico que estaba implícito ya en la obra de Löwenheim (1915), pero que la escuela de Hilbert tendía a desvirtuar.[1] El artículo es más conciso y elegante que la tesis y contiene un importante teorema que falta en ésta, pero omite casi todas las consideraciones metodológicas, no sé si en aras de la brevedad o para evitar polémicas. En la siguiente presentación utilizo ambas fuentes.[2]

El párrafo inicial del artículo plantea el problema con insuperable lucidez: En *Principia Mathematica*, Whitehead y Russell adoptan como axiomas ciertas proposiciones evidentes y proceden a deducir de ellas —conforme a ciertas reglas de inferencia precisas y sin prestar atención al significado de los símbolos— los teoremas de la lógica y las matemáticas. Ante este procedimiento, cabe preguntarse si el sistema de axiomas y reglas de inferencia propuesto es completo (*vollständig*), esto es, "si efectivamente basta

1 Piénsese que Herbrand llama *propositions vraies* justamente a las fórmulas *deducibles*.

2 La excelente versión castellana de las *Obras completas* de Gödel, editada por Jesús Mosterín, contiene el artículo (pp. 23–37), mas no la tesis. Ésta aparece, acompañada de traducción inglesa, en Gödel CW, vol. I, pp. 60–101.

para deducir cada teorema lógico-matemático, o si cabe tal vez concebir proposiciones verdaderas (y eventualmente también demostrables conforme a otros principios) que no se pueden derivar en el sistema en cuestión" (1930, p. 349).[3] Gödel recuerda que en el caso de las fórmulas verifuncionales esta cuestión había sido contestada afirmativamente por Bernays (1926) y anuncia que la resolverá, también afirmativamente, para una clase más amplia de fórmulas, a saber, las del cálculo predicativo de primer orden.[4]

La cuestión planteada supone que en el estudio de un cálculo lógico distingamos dos tipos de conceptos: (i) aquellos que "conciernen únicamente a los signos considerados como figuras espaciales" y (ii) aquellos cuya definición "tiene en cuenta el *significado* de las fórmulas" (Gödel 1929, § 2, en

3 Gödel 1929 declara, de entrada, que su propósito es probar que el "sistema axiomático del llamado cálculo funcional restringido" es completo, y explica que esta propiedad significa "que cada fórmula válida (*allgemein giltige*) expresable en [este cálculo] se puede deducir de los axiomas mediante una serie finita de inferencias formales" (CW, I, 60). En suma, un cálculo lógico premunido de axiomas y reglas de inferencia es *completo* en el sentido de Gödel 1929, 1930, si es *suficiente* para deducir todas las verdades en su campo de aplicación. Post 1921 estableció que el sistema deductivo del cálculo proposicional es suficiente o completo en un sentido puramente sintáctico: si $\phi$ es una fórmula del cálculo que no se puede deducir de los axiomas, entonces puede deducirse de ellos su negación $\neg\phi$. Tras observar que un sistema deductivo consistente para el cálculo predicativo de primer orden no podría ser completo en el sentido sintáctico, Hilbert y Ackermann (1928, p. 68) plantearon —quizás por vez primera— la cuestión de suficiencia abordada por Gödel. Dicen allí que "aún no se ha resuelto la cuestión de si el sistema axiomático [propuesto por ellos para ese cálculo] es completo en el sentido de que realmente se puedan derivar de él todas las fórmulas lógicas que son correctas (*richtig*) para todo dominio de individuos". Pero la expresión 'fórmulas lógicas' se introduce en ese libro de una manera que hace completamente incomprensible el planteamiento citado. Cito: "Entre las fórmulas que pueden demostrarse mediante el cálculo predicativo se destacan aquellas que no contienen constantes individuales y cuya derivación no presupone otras fórmulas que las fórmulas lógicas básicas [los axiomas]. Llamaremos a estas fórmulas *fórmulas lógicas*" (Hilbert y Ackermann 1928, p. 54). Según esta estipulación, pues, una fórmula lógica es *siempre* derivable y el cálculo es completo *por definición*.

4 Siguiendo a Hilbert y Ackermann (1928), Gödel lo llama *engeres Funktionenkalkül* (literalmente: 'más estrecho cálculo de funciones'). En una nota explica que pertenecen a él todas las "expresiones lógicas" formadas con variables proposicionales y funcionales —o sea, en nuestra jerga, con predicados $n$-ádicos, para cualquier $n \geq 0$— mediante las "operaciones" de disyunción, negación, cuantificación universal y cuantificación existencial, "en las cuales los prefijos $\forall x$ y $\exists x$ se refieren *sólo* a individuos, *no* a predicados (*Funktionen*)" (Gödel 1930, p. 349, n. 3).

CW, I, 66). Hoy los llamamos conceptos *sintácticos* y *semánticos*, respectivamente. Gödel va a probar que, en el cálculo predicativo de primer orden de *Principia Mathematica*, la clase de las fórmulas *deducibles* (propiedad *sintáctica*) coincide con la clase de las fórmulas que expresan *verdades lógicas* (propiedad *semántica*). Para ello, caracteriza primero el cálculo que va a considerar y da una definición de estos dos conceptos, aplicable a las fórmulas de ese cálculo.

El cálculo considerado por Gödel casi no se distingue del **CP1=**. Usaré, pues, nuestro simbolismo en lugar del suyo.[5] Con todo, seguiré su sintaxis en un respecto que no tiene su contraparte en la nuestra: Una letra mayúscula seguida de variables individuales entre paréntesis y separadas por comas representa una fórmula —normalmente, sin cuantificadores— en que las variables en cuestión ocupan posiciones libres. Según esto, la fórmula $F(x,y)$ puede corresponder a nuestra fórmula elemental $P^2xy$, pero también a una matriz compuesta como $\neg(P^2xy \to (P^1{}_1y \vee P^1{}_2x))$.[6] Al igual que Gödel, uso cualquier letra minúscula como variable individual y cualquier mayúscula como variable proposicional o en la función recién descrita. Gödel elige como *axiomas lógicos* las ocho fórmulas siguientes:

| | | | |
|---|---|---|---|
| **1** | $X \vee X \to X$ | **2** | $X \to X \vee Y$ |
| **3** | $X \vee Y \to Y \vee X$ | **4** | $(X \to Y) \to (Z \vee X \to Z \vee Y)$ |
| **5** | $\forall x F(x) \to F(y)$ | **6** | $\forall x(X \vee F(x)) \to X \vee \forall x F(x)$ |
| **7** | $x = x$ | **8** | $x = y \to (F(x) \to F(y))$ |

Las reglas de inferencia son: (I) *Modus ponens*. (II) “La regla de sustitución

[5] Gödel escribe & en vez de $\wedge$, $(x)$ en vez de $\forall x$ y $(Ex)$ en vez de $\exists x$. Para facilitar la lectura, utiliza paréntesis de diversas formas. Simboliza la negación, a la manera de Hilbert y Ackermann, mediante una línea recta trazada encima de todo su alcance. Así, por ejemplo, su fórmula $\overline{(x)(Ey)P(x,y)}$ corresponde a nuestra $\neg\forall x\exists yPxy$.

[6] Este simbolismo puede también interpretarse como una expresión del hecho siguiente: dada una formula con $n$ variables libres, cabe siempre *introducir* un predicado $n$-ario que la represente. Así, en el caso del ejemplo, se puede introducir un predicado binario $F$ definido por la equivalencia: $Fxy \leftrightarrow \neg(P^2xy \to (P^1{}_1y \vee P^1{}_2x))$.

para variables proposicionales y funcionales".[7] (III) De $A(x)$ se puede inferir $\forall xA(x)$ (generalización universal). (IV) Todas las variables —libres y ligadas— de una fórmula se pueden reemplazar por otras, siempre que se observen ciertas cautelas (dirigidas a impedir que la reemplazante de una variable libre sea "cazada" indebidamente por un cuantificador).[8] Hay que entender que '$(X \rightarrow Y)$' abrevia a '$(\neg X \vee Y)$' y que el cuantificador existencial y los conectivos $\wedge$ y $\leftrightarrow$ se definen en la forma habitual (Apéndice IX.D). Una fórmula $\phi$ es *deducible* (abreviado $\vdash\phi$) si hay una prueba de $\phi$, *refutable* si hay una prueba de $\neg\phi$. Gödel no explica el término 'prueba', pero evidentemente podemos adoptar una definición por el estilo de la que presenté en la p. 247: una lista de fórmulas $\langle\alpha_1, \ldots, \alpha_n\rangle$ es una prueba de $\alpha_n$ si y sólo si cada fórmula $\alpha_k$ es idéntica a uno de los ocho axiomas o se infiere de una fórmula $\alpha_j$ por las reglas II–IV, o de un par de fórmulas $\alpha_j$ y $\alpha_h$ por la regla I ($1 \leq k \leq n$; $j < h < k$).

Gödel explica así los conceptos semánticos esenciales de *validez* y *realizabilidad* (en la tesis; la explicación falta en el artículo). Sea $\alpha$ una fórmula que contiene las variables proposicionales $X_1, \ldots, X_m$, los predicados monádicos o poliádicos $F_1, \ldots, F_k$, las variables individuales libres $x_1, \ldots, x_l$, "y fuera de eso sólo variables ligadas" (CW, I, 66). Consideremos ahora ciertos "individuos" $a_1, \ldots, a_l$ pertenecientes a un cierto dominio (*Denkbereich*) no vacío $\mathfrak{D}$, ciertas "funciones (definidas en el mismo dominio)" $f_1, \ldots, f_k$ y aseveraciones (*Aussagen*) $A_1, \ldots, A_m$. "Diremos que este sistema $S = (f_1, \ldots, f_k, a_1, \ldots, a_l, A_1, \ldots, A_m)$ *realiza* (*erfülle*) la fórmula si, insertado (*eingesetzt*) en ella, genera una oración verdadera (en el dominio en cuestión)." Mediante esta definición, Gödel define los términos *realizable* (*erfüllbar*) *en un dominio determinado*, *realizable* absolutamente ("hay un dominio en el que la fórmula es realizable"), válido (*allgemein giltig*) en un dominio determinado ("su negación no es realizable") y *válido* absolutamente.

[7] Gödel no explica esta regla, pero entiendo que ella autoriza a reemplazar uniformemente (i) una variable proposicional por una fórmula cualquiera (siempre que ésta no contenga en posiciones libres una variable ligada por un cuantificador en cuyo alcance se hallaba alguna de las posiciones de la variable proposicional reemplazada) y (ii) cualquier subfórmula formada por un predicado $n$-ádico seguido de $n$ variables libres por una fórmula cualquiera cuyas variables libres sean precisamente las mismas.

[8] Gödel no detalla estas cautelas, sino que remite a Hilbert y Ackermann 1928, III, § 5.

Las fórmulas absolutamente válidas son las verdades lógicas. Las definiciones semánticas de Gödel no son impecables, pero apuntan sin duda a las que, siguiendo a Tarski, doy en el Apéndice IX.[9]

Es fácil comprobar que los axiomas **1**–**8** son válidos en el sentido indicado y que las reglas de inferencia I–IV preservan la validez. Gödel lo acepta sin discusión. Por lo tanto, *toda fórmula deducible es válida*. Lo que va a probar es la proposición recíproca: *Toda fórmula válida* del **CP1=** *es deducible* en el sistema deductivo descrito (Teorema I del artículo). Ella equivale a esta otra: *Toda fórmula* del **CP1=** *es realizable o refutable* (Teorema II). En efecto, si vale el Teorema I y la fórmula $\phi$ no es realizable, su negación $\neg\phi$ es válida y por ende deducible, de suerte que $\phi$ es refutable, como dice el Teorema II; mientras que si vale el Teorema II y $\phi$ es válida, $\neg\phi$ no es realizable y por ende es refutable, de suerte que $\neg\neg\phi$ es deducible y, con ella, también $\phi$, como dice el Teorema I. En vista de ello, Gödel acomete directamente la prueba del Teorema II.[10] Lo prueba primero para las fórmu-

9 Traduje lo más literalmente que pude el pasaje clave para criticar aquí sus defectos. Ante todo, ¿en qué consiste la operación de *insertar* (*einsetzen*) el sistema $S$ en la fórmula $A$? $S$ consta de elementos lingüísticos —las aseveraciones $A_1, \ldots, A_m$— que podrían, tal vez, insertarse literalmente en $A$, pero también incluye elementos no lingüísticos, que no sería posible acomodar dentro de una fórmula. Es razonable suponer que cuando habla de 'insertar' ciertos individuos en la fórmula Gödel quiere referirse al reemplazo uniforme de cada variable individual por una *expresión* que designe exclusivamente a uno de esos individuos. Otro tanto habría que decir de la inserción de las funciones $f_1, \ldots, f_k$. Hay que advertir, además, que cada una de éstas tendrá que ser lo que solía llamarse una "función proposicional", esto es, un aplicación de $\mathscr{D}^n$ en {0,1} o en {"lo verdadero", "lo falso"}, con $n$ igual a la $n$-adicidad del predicado de la lista $\langle F_1, \ldots, F_k\rangle$ en cuyo lugar la función se "inserta". En Gödel 1929 y 1930 falta toda indicación sobre el empleo de functores (signos que se combinan con términos para formar nuevos términos; *vide* Apéndice IX.E). Por otra parte, en vista del papel que desempeñan las variables proposicionales en el cálculo lógico, no hace falta, para juzgar la realizabilidad de una fórmula que contenga tales variables, que éstas se sustituyan con aseveraciones bien determinadas, como propone Gödel; basta asignarle a cada variable proposicional un valor veritativo fijo cualquiera. Finalmente, Gödel nos deja en la total oscuridad en lo que respecta a verdad o falsedad de la oración obtenida "insertando" el sistema $S$ en la fórmula $A$ cuando ésta contiene cuantificadores. Será Tarski quien, junto con elucidar la "inserción" como *interpretación*, enuncie con precisión las condiciones en que una fórmula precedida por $\forall$ o $\exists$ es o no verdadera en una interpretación dada (cf. las reglas [$\mathrm{I}_\forall$] e [$\mathrm{I}_\exists$] en el Apéndice IX.D).

10 El Teorema II de Gödel 1930 incluye también la siguiente cláusula: Si una fórmula no es refutable, es realizable en un dominio infinito enumerable. Con este agregado, el

las que no contienen el signo de identidad y cuya deducibilidad o refutabilidad depende sólo de los axiomas **1–6**, es decir, para el cálculo correspondiente a nuestro **CP1**. Luego, mediante un argumento sencillo, extiende el resultado al **CP1=**.

Sea $\mathfrak{F}$ la clase de las fórmulas del **CP1**, $\mathfrak{W}$ la subclase de las fórmulas refutables (mediante los axiomas **1–6** y las reglas I–IV) y $\mathfrak{R}$ la subclase de las fórmulas realizables. Se trata de probar que $\mathfrak{F} = \mathfrak{R} \cup \mathfrak{W}$. Para abreviar, llamaré $\mathfrak{T}$ a la clase $\mathfrak{R} \cup \mathfrak{W}$. En virtud del resultado de Bernays aludido arriba, es claro que toda fórmula verifuncional, esto es, toda fórmula que consta sólo de variables proposicionales, conectivos y paréntesis, pertenece a $\mathfrak{T}$. La discusión puede, entonces confinarse a las fórmulas del **CP1** que contienen también otros signos. Sea pues $\phi$ una fórmula cualquiera de este tipo. Gödel prueba primero que $\phi \in \mathfrak{T}$ si y sólo cierta fórmula de una clase especial $\mathfrak{K} \subseteq \mathfrak{F}$ pertenece a $\mathfrak{T}$. La etapa decisiva de la prueba consiste entonces en establecer que $\mathfrak{K} \subseteq \mathfrak{T}$. La reducción del problema general a este otro más particular comprende varios pasos que describo enseguida (en las notas doy una idea de la argumentación). Digamos, como de costumbre, que una fórmula $\psi$ equivale lógicamente a otra fórmula $\chi$ si la fórmula $(\psi \leftrightarrow \chi)$ es válida. Diré, por otra parte, que $\psi$ equivale *deductivamente* a $\chi$ si la fórmula $(\psi \leftrightarrow \chi)$ es deducible. Sin duda, hay una fórmula prenexa $\phi_1$ que equivale lógica y deductivamente a nuestra formula $\phi$.[11] Obviamente, $\phi \in \mathfrak{T}$ si y sólo si $\phi_1 \in \mathfrak{T}$. Gödel muestra sin dificultad que la fórmula prenexa $\phi_1 \in \mathfrak{T}$

---

Teorema II es estrictamente más fuerte que el Teorema I y contiene (una forma de) el Teorema de Löwenheim-Skolem. La prueba gödeliana nos rinde pues este importante teorema como suplemento. (La cláusula mencionada se aplica a las fórmulas del **CP1**; como se verá en la nota 27, en el caso del **CP1=** vale la condición más débil: Si una fórmula no es refutable, es realizable en un dominio enumerable, que puede ser finito. Piénsese que la fórmula $\forall x \forall y(x = y)$, aunque irrefutable en el sistema descrito, únicamente es realizable en dominios que contengan un solo objeto).

11 En efecto, las fórmulas representadas por los esquemas (2)–(5) y (8)–(14) del Apéndice XIII son todas válidas y se deducen de los axiomas, usando las definiciones de los símbolos lógicos. En particular, el esquema (2) *es* la definición del cuantificador existencial; (3)–(5) se deducen fácilmente de (2). La demostración de los esquemas (8)–(14) depende decisivamente del axioma **6**.

12 Recuérdese que llamamos *cerrada* a una fórmula que no contiene variables libres. Si la fórmula prenexa $\phi_1$ contiene las variables libres $\xi_1, \xi_2, \ldots, \xi_k$ es claro que $\phi_1$ es realizable si y sólo si es realizable $\phi_2 = \exists\xi_1\exists\xi_2 \ldots \exists\xi_k\phi_1$. Tenemos además que, si $\neg\phi_1$

si y sólo si hay una fórmula prenexa cerrada $\phi_2 \in \mathfrak{T}$.[12] Por otra parte, la fórmula prenexa cerrada $\phi_2 \in \mathfrak{T}$ si y sólo si hay una fórmula prenexa cerrada $\phi_3$, cuyo prefijo comienza con un cuantificador universal y termina con un cuantificador existencial, tal que $\phi_3 \in \mathfrak{T}$.[13] Las fórmulas prenexas cerradas cuyo prefijo comienza con un cuantificador universal y termina con un cuantificador existencial constituyen la clase especial $\mathfrak{K}$. Llamémoslas $\mathfrak{K}$-*fórmulas*. El *grado* de una $\mathfrak{K}$-fórmula es el número de cuantificadores universales en su prefijo cuyos respectivos sucesores *inmediatos* son cuantificadores existenciales. Sea $\mathfrak{K}_n$ la clase de las $\mathfrak{K}$-fórmulas de grado *n*. Gödel

---

es deducible, la regla III permite deducir $\forall\xi_1\forall\xi_2 \ldots \forall\xi_k\neg\phi_1$ y, por consiguiente, $\neg\exists\xi_1\exists\xi_2 \ldots \exists\xi_k\phi_1$; y que, si $\neg\exists\xi_1\exists\xi_2 \ldots \exists\xi_k\phi_1$ es deducible, también es deducible $\forall\xi_1\forall\xi_2 \ldots \forall\xi_k\neg\phi_1$, de la cual se deduce $\neg\phi_1$ por aplicaciones repetidas del axioma **5** y *modus ponens*. Por lo tanto, $\phi_1$ es refutable si y sólo si lo es $\phi_2$.

13 Sea $\phi_2$ la fórmula prenexa cerrada $\mathbf{Q\xi}\psi$, donde $\psi$ es una matriz y el prefijo $\mathbf{Q\xi}$ es una fila formada por uno o más cuantificadores de cualquier clase. Considérese la fórmula $\phi_3 = \forall\zeta\mathbf{Q\xi}\exists\eta(\psi \wedge (\Phi\zeta \vee \neg\Phi\eta))$, donde $\Phi$ es un predicado monádico y $\zeta$ y $\eta$ son variables individuales que no figuran en $\psi$ (como hemos visto en otras ocasiones, esta indicación puede precisarse apelando al orden lexicográfico de predicados y variables: úsense *los primeros* que no figuran en $\psi$). $\phi_3$ es lógica y deductivamente equivalente a $\mathbf{Q\xi}\psi \wedge \forall\zeta\exists\eta(\Phi\zeta \vee \neg\Phi\eta)$. Como $\forall\zeta\exists\eta(\Phi\zeta \vee \neg\Phi\eta)$ es a la vez deducible y válida, $\phi_3$ es lógica y deductivamente equivalente a $\mathbf{Q\xi}\psi = \phi_2$, de modo que $\phi_2$ sólo será realizable o refutable si $\phi_3$ lo es.

14 Sea $\psi \in \mathfrak{K}_{n+1}$. Si $n > 1$, $\psi$ tiene la forma $\mathbf{\forall\xi\exists\zeta\forall\chi\exists\eta Q\upsilon}\Psi$, donde $\Psi$ es una matriz, $\mathbf{Q\upsilon}$ es una fila de cuantificadores de cualquier clase que comienza con uno universal y termina con uno existencial, y $\mathbf{\forall\xi}$, $\mathbf{\exists\zeta}$, $\mathbf{\forall\chi}$ y $\mathbf{\exists\eta}$ son filas de cuantificadores de la clase indicada. Obviamente, $\mathbf{\forall\chi\exists\eta Q\upsilon}\Psi \in \mathfrak{K}_n$ y $\mathbf{Q\upsilon}\Psi \in \mathfrak{K}_{n-1}$. Sean $\beta$ y $\gamma$ las fórmulas definidas como sigue:

$$\beta = \mathbf{\forall\xi'\exists\zeta'}\Phi(\xi',\zeta') \wedge \mathbf{\forall\xi\forall\zeta}(\Phi(\xi,\zeta) \to \mathbf{\forall\chi\exists\eta Q\upsilon}\Psi)$$

$$\gamma = \mathbf{\forall\xi'\forall\xi\forall\zeta\forall\chi\exists\zeta'\exists\eta Q\upsilon}(\Phi(\xi',\zeta') \wedge (\Phi(\xi,\zeta) \to \Psi))$$

No es difícil ver que $\vdash\beta \to \psi$. Invocando dos veces el teorema **B** enunciado más adelante en el texto (en la forma más general explicada entre corchetes), combinado con el teorema **C**, el lector podrá establecer que $\vdash\beta \leftrightarrow \gamma$. Ahora bien, $\gamma \in \mathfrak{K}_n$, de modo que, por hipótesis $\gamma \in \mathfrak{T}$. Si $\gamma$ es realizable, $\beta$ es realizable y por ende también lo es $\psi$. Si $\gamma$ es refutable, $\beta$ es refutable, de modo que $\vdash\neg\beta$. Reemplazando en $\beta$, $\Phi$ por $\mathbf{\forall\chi\exists\eta Q\upsilon}\Psi$ tenemos que

$$\vdash\neg(\mathbf{\forall\xi'\exists\zeta'\forall\chi\exists\eta Q\upsilon}\Psi \wedge \mathbf{\forall\xi\forall\zeta}(\mathbf{\forall\chi\exists\eta Q\upsilon}\Psi \to \mathbf{\forall\chi\exists\eta Q\upsilon}\Psi))$$

La subfórmula a la derecha del signo $\wedge$ es, por cierto, deducible. Por lo tanto, $\vdash\neg\mathbf{\forall\xi'\exists\zeta'\forall\chi\exists\eta Q\upsilon}\Psi$. Mediante una sustitución juiciosa de variables ligadas en la fór-

prueba que si $\mathfrak{K}_n \subseteq \mathfrak{T}$, entonces $\mathfrak{K}_{n+1} \subseteq \mathfrak{T}$.[14] Por lo tanto, si podemos probar que $\mathfrak{K}_1 \subseteq \mathfrak{T}$, concluiremos, por inducción sobre el grado de las $\mathfrak{K}$-fórmulas, que $\mathfrak{K} \subseteq \mathfrak{T}$, y por ende —en virtud de los pasos anteriores— que $\mathfrak{F} = \mathfrak{T}$. Como la clase $\mathfrak{T}$ reúne las fórmulas refutables y realizables, ella es un híbrido sintáctico-semántico y la prueba de que $\mathfrak{K}_1$ está incluida en $\mathfrak{T}$ tiene que participar de ambos aspectos. En su aspecto sintáctico, es larga y tediosa, pero elemental; en su aspecto semántico, es breve y brillante, pero nada elemental.

Consideramos una $\mathfrak{K}$-fórmula cualquiera $\psi$ de grado 1 en la que figuran $r$ variables ligadas por cuantificadores universales y $s$ variables ligadas por cuantificadores existenciales. $\psi$ tiene la forma $\boldsymbol{\forall\xi\exists\zeta}\Theta(\xi_1,\ldots,\xi_r,\zeta_1,\ldots,\zeta_s)$, donde $\boldsymbol{\forall\xi}$ y $\boldsymbol{\exists\zeta}$ son filas de cuantificadores de la clase indicada, que ligan las variables $\xi_1,\ldots, \xi_r$ y $\zeta_1,\ldots, \zeta_s$, respectivamente, y $\Theta(\xi_1,\ldots,\xi_r,\zeta_1,\ldots,\zeta_s)$ es una matriz en que figuran las variables listadas (y sólo ellas). Con una maniobra similar a la ejecutada contemporáneamente por Herbrand (Capítulo 2.6), Gödel asocia a la fórmula $\psi \in \mathfrak{K}_1$ una secuencia de matrices $\Theta_1, \Theta_2, \ldots$ construidas como ahora explicaré. Cada variable individual del cálculo consta de una letra $x$ seguida de $n$ palotes ($n \geq 0$). Forman, pues, un conjunto infinito numerable $\mathcal{V} = \{x_0, x_1, \ldots\}$, ordenado por el número de palotes de cada variable. Consideremos la familia de todos los $r$-tuplos que pueden formarse con elementos tomados de $\mathcal{V}$. Esta familia se puede ordenar atendiendo al número total de palotes contenido en cada $r$-tuplo; si dos $r$-tuplos contienen el mismo número de palotes, los ordenamos en orden lexicográfico. Así se forma la secuencia de $r$-tuplos $\mathbf{x}_1 = \langle x_0,x_0,\ldots,x_0,x_0\rangle$, $\mathbf{x}_2 = \langle x_0,x_0,\ldots,x_0,x_1\rangle$, $\mathbf{x}_3 = \langle x_0,x_0,\ldots,x_1,x_0\rangle, \ldots$ Las matrices $\Theta_1, \Theta_2, \ldots$ se forman sustituyendo uniforme y ordenadamente en $\Theta(\xi_1,\ldots,\xi_r,\zeta_1,\ldots,\zeta_s)$ las variables listadas por diversos $(r+s)$-tuplos tomados de $\mathcal{V}$. En cada caso, sustituimos el $r$-tuplo $\langle\xi_1,\ldots,\xi_r\rangle$ por uno de los $r$-tuplos de nuestra secuencia $\mathbf{x}_1, \mathbf{x}_2, \ldots$ y el $s$-tuplo $\langle\zeta_1,\ldots,\zeta_s\rangle$ por las primeras $s$ variables de $\mathcal{V}$ que no se hayan utilizado todavía como sustitutos. Específicamente,

---

mula precedente, obtenemos $\vdash\neg\boldsymbol{\forall\xi\exists\zeta\forall\chi\exists\eta Q\upsilon}\Psi$, o sea, $\vdash\neg\psi$. Así pues, $\psi$ es refutable si $\gamma$ lo es. Por consiguiente, $\psi \in \mathfrak{T}$. Para obtener la misma conclusión en el caso $n = 1$ basta eliminar del razonamiento anterior la expresión '$\mathbf{Q\upsilon}$' y su descripción.

$\Theta_1 = \Theta(\mathbf{x}_1, x_1, x_2, \ldots, x_s)$
$\Theta_2 = \Theta_1 \wedge \Theta(\mathbf{x}_2, x_{s+1}, x_{s+2}, \ldots, x_{2s})$
$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$
$\Theta_n = \Theta_{n-1} \wedge \Theta(\mathbf{x}_n, x_{(n-1)s+1}, x_{(n-1)s+2}, \ldots, x_{ns})$
$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$

Siguiendo a Gödel, voy a usar varias abreviaturas. Llamo $(P_n)$ al prefijo $\exists x_0 \exists x_1 \ldots \exists x_{ns}$. Recordando como está construida la secuencia $\mathbf{x}_1$, $\mathbf{x}_2$, . . . , vemos que el prefijo $(P_n)$ inevitablemente incluirá todas las variables del $r$-tuplo $\mathbf{x}_{n+1}$. Llamo $(P_n')$ al prefijo que se obtiene al eliminar de $(P_n)$ a todos los cuantificadores que ligan variables contenidas en $\mathbf{x}_{n+1}$. Por último, llamo $\mathbf{y}_n$ al $s$-tuplo de variables nuevas $\langle x_{(n-1)s+1}, x_{(n-1)s+2}, \ldots, x_{ns} \rangle$, que figuran en $\Theta_n$ pero no en $\Theta_{n-1}$. Conforme a esta convención escribimos, pues,

$$\Theta_n = \Theta_{n-1} \wedge \Theta(\mathbf{x}_n, \mathbf{y}_n) \tag{1}$$

El aspecto sintáctico de la prueba consiste principalmente en establecer que, para cada entero positivo $n$, la fórmula $(\psi \to (P_n)\Theta_n)$ es deducible. Sea $n = 1$. Por simple sustitución de variables ligadas, probamos que

$$\vdash \boldsymbol{\forall\xi\exists\zeta}\Theta(\xi_1, \ldots, \xi_r, \zeta_1, \ldots, \zeta_s) \to \boldsymbol{\forall}\mathbf{x}_1 \boldsymbol{\exists}\mathbf{y}_1 \Theta(\mathbf{x}_1, \mathbf{y}_1) \tag{2}$$

(donde $\boldsymbol{\forall}\mathbf{x}_1$ y $\boldsymbol{\exists}\mathbf{y}_1$ son filas de cuantificadores que ligan, respectivamente, a las variables de $\mathbf{x}_1$ e $\mathbf{y}_1$).[15] También se muestra fácilmente que

$$\vdash \boldsymbol{\forall}\mathbf{x}_1 \boldsymbol{\exists}\mathbf{y}_1 \Theta(\mathbf{x}_1, \mathbf{y}_1) \to \boldsymbol{\exists}\mathbf{x}_1 \boldsymbol{\exists}\mathbf{y}_1 \Theta(\mathbf{x}_1, \mathbf{y}_1) \tag{3}$$

Por la transitividad de la flecha, de (1) y (2) se infiere $\vdash \psi \to (P_1)\Theta_1$. La tesis general quedará establecida inductivamente si podemos mostrar que, para cada $n \geq 1$, $\vdash(\psi \wedge (P_n)\Theta_n \to (P_{n+1})\Theta_{n+1})$. Como en el caso de (1), es claro que

15 Como los $r$ elementos de $\mathbf{x}_1$ son la misma variable $x_0$, sólo el último cuantificador de la fila $\boldsymbol{\forall}\mathbf{x}_1$ liga en efecto a esa variable en todas sus posiciones libres en la matriz $\Theta(\mathbf{x}_1, x_1, x_2, \ldots, x_s)$; los $r$–1 restantes podrían borrarse. Pero evidentemente es más cómodo escribirlos (abreviadamente), lo que ciertamente está permitido por la sintaxis del **CP1** (Apéndice IX.A).

$$\vdash \forall\boldsymbol{\xi}\exists\boldsymbol{\zeta}\Theta(\xi_1,\ldots,\xi_r,\zeta_1,\ldots,\zeta_s) \rightarrow \forall\mathbf{x}_{n+1}\exists\mathbf{y}_{n+1}\Theta(\mathbf{x}_{n+1},\mathbf{y}_{n+1}) \quad (4)$$

Permutando los cuantificadores existenciales que forman el prefijo $(P_n)$, obtenemos

$$\vdash((P_n)\Theta_n \rightarrow \exists\mathbf{x}_{n+1}(P_n')\Theta_n) \quad (5)$$

Para completar la prueba, Gödel invoca tres teoremas familiares de la lógica de primer orden (el lector que no los conozca puede ejercitarse en probar al menos los dos primeros):[16]

**A** Si $\boldsymbol{\xi}$ es un *r*-tuplo de variables y $\alpha(\boldsymbol{\xi})$ y $\beta(\boldsymbol{\xi})$ son dos fórmulas en que esas variables ocupan todas las posiciones libres, entonces:

$$\vdash\forall\boldsymbol{\xi}\alpha(\boldsymbol{\xi}) \wedge \exists\boldsymbol{\xi}\beta(\boldsymbol{\xi}) \rightarrow \exists\boldsymbol{\xi}(\beta(\boldsymbol{\xi}) \wedge \alpha(\boldsymbol{\xi})).$$

**B** Sea $\mathbf{Q}\boldsymbol{\xi}$ una fila homogénea de cuantificadores (todos existenciales o todos universales) y $\mathbf{Q}\boldsymbol{\zeta}$ otra fila, homogénea también. Sea $(\Pi)$ el prefijo que resulta al combinar en cualquier orden los cuantificadores de ambas filas. Entonces, si $\alpha(\boldsymbol{\xi})$ y $\beta(\boldsymbol{\zeta})$ son dos fórmulas en que las variables listadas ocupan posiciones libres,

$$\vdash\mathbf{Q}\boldsymbol{\xi}\alpha(\boldsymbol{\xi}) \wedge \mathbf{Q}\boldsymbol{\zeta}\beta(\boldsymbol{\zeta}) \leftrightarrow (\Pi)(\alpha(\boldsymbol{\xi}) \wedge \beta(\boldsymbol{\zeta}))$$

[Si $\mathbf{Q}\boldsymbol{\xi}$, o $\mathbf{Q}\boldsymbol{\zeta}$, o ambos, son filas de cuantificadores no homogéneas, lo dicho vale siempre que los cuantificadores provenientes de cada prefijo conserven en $(\Pi)$ el mismo orden que tenían en el prefijo de origen.]

**C** Si la fórmula $\alpha$ figura una o más veces como subfórmula en la

[16] Los teoremas **A**, **B** y **C** corresponden, respectivamente, a los "teoremas auxiliares" 1b, 4 y 6, enunciados al comienzo de Gödel 1930. El teorema auxiliar 4 expresa el caso general de **B**, que formulo entre corchetes al final,. He preferido destacar el caso especial en que todos los cuantificadores son de la misma clase, porque ese es el que aplicaremos luego. El caso general se emplea en el argumento esbozado en la nota 14.

fórmula $\theta(\alpha)$ y $\theta(\beta)$ es la fórmula obtenida al reemplazar $\alpha$ por la fórmula $\beta$, deductivamente equivalente a $\alpha$, en una o más de las posiciones de $\alpha$ en $\theta(\alpha)$, entonces $\theta(\alpha)$ y $\theta(\beta)$ son deductivamente equivalentes. En otras palabras, $\vdash(\alpha \leftrightarrow \beta)$ implica que $\vdash(\theta(\alpha) \leftrightarrow \theta(\beta))$.[17]

En la fórmula destacada en el teorema **A** sustituimos la subfórmula $\alpha(\boldsymbol{\xi})$ por $\exists \mathbf{y}_{n+1}\Theta(\mathbf{x}_{n+1},\mathbf{y}_{n+1})$ y la subfórmula $\beta(\boldsymbol{\xi})$ por $(P_n{}')\Theta_n$ (donde, como se hizo notar, las $r$ variables contenidas en $\mathbf{x}_{n+1}$ ocupan posiciones libres); sustituimos, asimismo, en los prefijos $\forall\boldsymbol{\xi}$ y $\exists\boldsymbol{\xi}$ el $r$-tuplo $\boldsymbol{\xi}$ por $\mathbf{x}_{n+1}$. Resulta, entonces, que

$$\begin{aligned}\vdash\forall\mathbf{x}_{n+1}\exists\mathbf{y}_{n+1}\Theta(\mathbf{x}_{n+1},\mathbf{y}_{n+1}) \wedge \exists\mathbf{x}_{n+1}(P_n{}')\Theta_n \\ \rightarrow \exists\mathbf{x}_{n+1}((P_n{}')\Theta_n \wedge \exists\mathbf{y}_{n+1}\Theta(\mathbf{x}_{n+1},\mathbf{y}_{n+1}))\end{aligned} \qquad (6)$$

Ahora bien, las apódosis de (4) y (5) —esto es, las subfórmulas que siguen a la flecha en dichas fórmulas— son los dos miembros de la conjunción que forma la prótasis —la subfórmula que precede a la flecha— en (6). Podemos, pues, reemplazar ésta por la conjunción de las prótasis de (4) y (5),[18] para obtener:

$$\vdash\psi \wedge (P_n)\Theta_n \rightarrow \exists\mathbf{x}_{n+1}((P_n{}')\Theta_n \wedge \exists\mathbf{y}_{n+1}\Theta(\mathbf{x}_{n+1},\mathbf{y}_{n+1})) \qquad (7)$$

En la fórmula destacada en el teorema **B** ponemos $n = (n + 1)s$ y $m = (ns - r)$ y sustituimos $\mathbf{Q}\boldsymbol{\xi}^n\alpha(\boldsymbol{\xi}^n)$ por $\exists\mathbf{y}_{n+1}\Theta(\mathbf{x}_{n+1},\mathbf{y}_{n+1})$ y $\mathbf{Q}\boldsymbol{\zeta}^m\beta(\boldsymbol{\zeta}^m)$ por $(P_n{}')\Theta_n$. Como el prefijo $(P_{n+1}{}')$ se obtiene uniendo y reordenando los cuantificadores —existenciales todos— en $\exists\mathbf{y}_{n+1}$ y $(P_n{}')$, podemos identificarlo con el prefijo $(\Pi)$. Tenemos, entonces, que

$$\vdash(P_n{}')\Theta_n \wedge \exists\mathbf{y}_{n+1}\Theta(\mathbf{x}_{n+1},\mathbf{y}_{n+1}) \leftrightarrow (P_{n+1}{}')(\Theta_n \wedge \Theta(\mathbf{x}_{n+1},\mathbf{y}_{n+1})) \qquad (8)$$

Pero $\Theta_n \wedge \Theta(\mathbf{x}_{n+1},\mathbf{y}_{n+1}) = \Theta_{n+1}$. Por lo tanto, (8) y **B** juntos implican que

17 El "metateorema" **C** refleja el carácter "extensional" propio del cálculo predicativo. Todo buen manual de lógica matemática lo demuestra para el sistema deductivo adoptado en él. Gödel 1930, p. 351, remite a Hilbert y Ackermann 1928, III, § 7.

18 Puesto que $\vdash(A \rightarrow B) \wedge (C \rightarrow D) \rightarrow ((B \wedge D \rightarrow F) \rightarrow (A \wedge C \rightarrow F))$.

$$\vdash \mathbf{\exists x}_{n+1}((P_n{}')\Theta_n \wedge \mathbf{\exists y}_{n+1}\Theta(\mathbf{x}_{n+1},\mathbf{y}_{n+1})) \leftrightarrow (P_{n+1})\Theta_{n+1} \qquad (9)$$

El lado izquierdo de la equivalencia (9) es la apódosis de (7). Reemplazándolo por el lado derecho, obtenemos el resultado que buscábamos:

$$\vdash \psi \wedge (P_n)\Theta_n \rightarrow (P_{n+1})\Theta_{n+1} \qquad (10)$$

Habiendo establecido que $\vdash \psi \rightarrow (P_1)\Theta_1$ y que, para cualquier $n > 0$, $\vdash \psi \wedge (P_n)\Theta_n \rightarrow (P_{n+1})\Theta_{n+1}$, podemos concluir que $\vdash \psi \rightarrow (P_n)\Theta_n$ para cada $n > 0$. Por contraposición, $\vdash \neg(P_n)\Theta_n \rightarrow \neg\psi$. Por lo tanto, si $(P_n)\Theta_n$ es refutable en el caso de algún entero positivo $n$, nuestra fórmula $\psi$, sujeta a la sola condición de ser una $\mathfrak{K}$-fórmula de grado 1, también es refutable. Ahora bien, $(P_n)\Theta_n$ es refutable si $\vdash \neg(P_n)\Theta_n$, o sea, si $\vdash \neg\Theta_n$ (en virtud de la regla III, puesto que $(P_n)$ es una fila de cuantificadores existenciales y $\vdash \forall\xi\neg\alpha \leftrightarrow \neg\exists\xi\alpha$). Como $\Theta_n$ es una matriz, hay una fórmula verifuncional $\vartheta_n$ que se forma reemplazando uniformemente cada subfórmula elemental de $\Theta_n$ (cada predicado $n$-ádico seguido de $n$ términos) por una variable proposicional (predicado 0-ádico) diferente. (Para mayor comodidad en la exposición, supongamos que cada variable proposicional de $\Theta_n$ es reemplazada en $\vartheta_n$ por ella misma.) Aplicando juiciosamente la regla de sustitución II, cualquier prueba de $\neg\vartheta_n$ se transforma en una prueba de $\neg\Theta_n$. Por lo tanto, $(P_n)\Theta_n$ es refutable si la fórmula verifuncional $\vartheta_n$ es refutable. Ahora bien, como sabemos, si $\vartheta_n$ *no es* refutable, $\vartheta_n$ es realizable, es decir, hay una asignación de valores veritativos a las variables proposicionales que figuran en $\vartheta_n$ que hace "verdadera" a $\vartheta_n$. En la decisiva fase semántica de la prueba, Gödel explota este hecho para establecer que $\psi$ es realizable cuando no es refutable, esto es, que $\psi \in \mathfrak{T}$.

Para ello, hay que atender a la estructura interna de $\psi$. Sabemos que contiene las variables $\xi_1, \ldots, \xi_r, \zeta_1, \ldots, \zeta_s$. A la luz de lo que ahora veremos,

[19] En otras palabras, la prueba de Gödel está confinada a lo que suele llamarse cálculo predicativo *puro* (sin constantes individuales). Pero el argumento sintáctico arriba desarrollado no supone esta restricción y el argumento semántico que ahora veremos puede adaptarse sin dificultad al caso en que la fórmula $\psi$ bajo consideración contiene constantes individuales (es cosa de darles también a éstas una interpretación numérica como la que se les da a las variables).

es claro que Gödel da por descontado que no contiene otros términos.[19] Supongamos que contiene las variables proposicionales $X_1, \ldots, X_h$, y los predicados ν-ádicos $F_1, \ldots, F_k$ (para diversos valores de $\nu > 0$). Como es obvio, $X_1, \ldots, X_h$ y $F_1, \ldots, F_k$ también están contenidos en cada matriz $\Theta_n$ y son todas las variables proposicionales y predicados ν-ádicos que allí figuran. Gödel nos invita a considerar, para un dado $n > 0$, un conjunto de $h$ valores veritativos $w_{1n}, \ldots, w_{hn}$ y $k$ funciones proposicionales $f_{1n}, \ldots, f_{kn}$, definidas en el conjunto de enteros $Z_n = \{z : 0 \leq z \leq ns\}$ y elegidas de modo que $f_{jn}$ tenga precisamente la misma ν-adicidad que $F_j$ $(1 \leq j \leq k)$.[20] Diremos, con Gödel, que este conjunto es un "sistema de realización de nivel $n$-ésimo" (*Erfüllungssystem n-ter Stufe*) de la fórmula ψ si es verdadera la aseveración que se obtiene a partir de la matriz $\Theta_n$ al asignar el valor veritativo $w_{in}$ a cada variable proposicional $X_i$ $(1 \leq i \leq h)$, interpretar cada predicado $F_j$ como representativo de la función proposicional $f_{jn}$ $(1 \leq j \leq k)$ y reemplazar cada una de las variables $x_0, \ldots, x_n$ que figuran en $\Theta_n$ por su respectivo índice numérico (contenido en $Z_n$). Si $S_m$ y $S_n$ son dos sistemas de realización de ψ de nivel $m$-ésimo y $n$-ésimo, respectivamente $(m < n)$, diremos que $S_n$ incluye a $S_m$ $(S_m \subset S_n)$ si cada función $f_{jm} \in S_m$ es la restricción a $Z_m$ de la función $f_{jn} \in S_n$ y los valores veritativos $w_{im}$ y $w_{in}$ son, respectivamente, iguales para cada índice $i$ $(1 \leq j \leq k,\ 1 \leq i \leq h)$.

Es claro que, si la fórmula verifuncional $\vartheta_n$ asociada a la matriz $\Theta_n$ es realizable, tiene que haber una interpretación del tipo descrito en que sea verdadera, esto es, un sistema de realización de ψ de nivel $n$-ésimo.[21] Como

20 Siguiendo a Gödel, digo simplemente que la función ν-ádica $f$ está definida en $Z_n$ si está definida en $(Z_n)^\nu$.

21 El sencillo truco siguiente genera la interpretación deseada, como el lector podrá fácilmente comprobar. Recordemos que $\vartheta_n$ se forma reemplazando cada subfórmula elemental de $\Theta_n$ por una variable proposicional. Supongamos, para mayor comodidad, que cada variable proposicional $X_i$ está representada en $\vartheta_n$ por ella misma. Las demás subfórmulas elementales de $\Theta_n$ pueden numerarse en el orden en que aparecen por primera vez. Sea $Y_q$ la variable proposicional que reemplaza a la $q$-ésima subfórmula. Si $\vartheta_n$ es realizable, hay una asignación $\mathcal{V}$ de valores veritativos a las $X_i$ y las $Y_q$ tal que $\mathcal{V}(\vartheta_n) = 0$ ("lo verdadero"). Ponemos $w_{in} = \mathcal{V}(X_i)$. Si $F_j$ es ν-ádico, definimos la función proposicional $f_{jn}$: $(Z_n)^\nu \rightarrow \{0,1\}$ como sigue: si hay un entero $q$ tal que la $q$-ésima subfórmula elemental de $\Theta_n$ está formada por el predicado $F_j$ seguido por ν variables con los índices numéricos $z_1, \ldots, z_\nu$, respectivamente, entonces $f_{jn}(z_1, \ldots, z_\nu) = \mathcal{V}(Y_q)$; en cualquier otro caso $f_{jn}(z_1, \ldots, z_\nu) = 0$.

hemos visto, si ψ no es refutable, $\vartheta_n$ es realizable y hay un sistema de realización de ψ de nivel $n$-ésimo para cada $n > 0$. Teniendo en cuenta que $\Theta_{n+1} = \Theta_n \wedge \Theta(\mathbf{x}_{n+1}, x_{ns+1}, x_{ns+2}, \ldots, x_{(n+1)s})$, es claro que cualquier sistema de realización de ψ de nivel ($n$+1)-ésimo provee una realización de $\Theta_n$ y por lo tanto incluye un sistema de realización de ψ de nivel $n$-ésimo.[22] Según Gödel, de esta sencilla observación "se sigue, con arreglo a formas de inferencia conocidas (*nach bekannten Schlußweisen*), que en tal caso hay una secuencia de sistemas de realización $S_1, S_2, \ldots, S_k, \ldots$ (con $S_k$ de nivel $k$-ésimo), en que cada uno de los que siguen al primero contiene al que le precede" (1930, p. 356). Este es el paso decisivo en el argumento semántico de Gödel, así que vale la pena que lo examinemos con atención. En la tesis, lo explica un poco más. Anota con razón que, para cada $n > 0$, el número de sistemas de realización de ψ de nivel $n$-ésimo tiene que ser finito (pues es finito el número total de ($h$+$k$)-tuplos diferentes formados por $h$ valores veritativos y $k$ funciones proposicionales definidas en $Z_n$). "Por lo tanto, al menos *uno* de los sistemas de realización de primer nivel tiene que estar contenido en infinitos otros de nivel superior. Hay, pues, un sistema de realización de segundo nivel en que *ese* sistema está contenido, el cual a su vez está contenido en infinitos otros de nivel superior. Continuando este razonamiento se muestra en forma conocida la existencia de una secuencia de sistemas de realización $S_1, S_2, \ldots, S_i, \ldots$, *ad infinitum*, donde $S_i \subset S_{i+1}$ y $S_i$ es de nivel $i$-ésimo" (CW, I, 86).[23]

¿Cuál es esa forma conocida de razonar a que se refiere Gödel? Contra lo que el pasaje anterior parecería destinado a sugerir, no puede tratarse de la inducción matemática. Ésta supone que, dado un objeto caracterizable con un índice numérico $n$, sepamos construir o señalar un objeto análogo con el índice $n + 1$. Pero en la situación que confrontamos aquí las cosas se presentan al revés: con cada sistema de realización de nivel $n > 1$ está dado otro de nivel $n$–1, contenido en el primero. Repitiendo el razonamiento, obtenemos una serie finita de sistemas de realización en forma de muñeca

22 Específicamente, si $\{w_1,\ldots,w_h,f_1,\ldots,f_k\}$ es un sistema de realización de ψ de nivel ($n$+1)-ésimo, y designamos con $f'_j$ a la restricción de $f_j$ a $Z_n$, $\{w_1,\ldots,w_h,f'_1,\ldots,f'_k\}$ es un sistema de realización de ψ de nivel $n$-ésimo incluido en $\{w_1,\ldots,w_h,f_1,\ldots,f_k\}$.

23 En el original, la última oración dice así : "So weiter schließend zeigt man in bekannter Weise die Existenz einer Folge von Erfüllungssystemen $S_1, S_2, \ldots, S_i, \ldots$, ad infinitum, wobei $S_i \subset S_{i+1}$ und $S_i$ von $i$-ter Stufe ist."

rusa: $S_n \supset S_{n-1} \supset S_{n-2} \supset \ldots \supset S_2 \supset S_1$; pero eso no prueba que dicha serie coincida con la parte final de la serie obtenida a partir de otro sistema $S_m$, con $m > n$. Por hipótesis, tenemos una galería infinita de "muñecas rusas" como la descrita, una para cada entero positivo $n$, pero no cabe combinarlas, por inducción, en una sola supermuñeca infinita. Por otra parte, hay una proposición —que Dénes König (1926) dedujo del Axioma de Selección, pero que es estrictamente más débil que éste— que viene como anillo al dedo a la situación descrita por Gödel. Dice así:

**K** Si $\Sigma_1, \Sigma_2, \ldots$ es una secuencia de conjuntos finitos no vacíos y hay una relación binaria $R$ tal que para cada $a \in \Sigma_{n+1}$ hay un $b \in \Sigma_n$ tal que $bRa$, entonces hay una secuencia $a_1, a_2, \ldots$ tal que, para cada $n > 0$, $a_n \in \Sigma_n$ y $a_nRa_{n+1}$.[24]

Para aplicarla a nuestro caso, basta identificar cada $\Sigma_n$ con el conjunto de todos los sistemas de realización de $\psi$ de nivel $n$-ésimo y escribir $\subset$ en vez de $R$. Como Gödel no menciona a König, sus editores sugieren que el pasaje citado no alude a la misma proposición **K**, sino más bien a los argumentos —pertenecientes al "folklore" matemático— que movieron a König a adop-

24 La siguiente proposición, equivalente a **K**, se suele llamar *Lema del Infinito de König* en la teoría de grafos: Si un grafo infinito $G$ contiene subgrafos $G_1, G_2, \ldots$, tales que ninguno de ellos tiene un punto en común con otro, pero cada punto de $G_{n+1}$ está unido a un punto de $G_n$, entonces existe un camino infinito $p_1, p_2, \ldots$, tal que $p_i \in G_i$.

25 También Hilbert y Bernays, cuando reformulan la prueba de Gödel en *Grundlagen der Mathematik*, se abstienen de aducir expresamente el Lema de König u otra consecuencia del Axioma de Selección, y presentan el paso al infinito en el argumento semántico como un simple ejercicio del sano sentido común. Vale la pena ver cómo lo hacen. En vez de nuestros sistemas de realización de $\psi$ (que son conjuntos finitos de valores veritativos y funciones proposicionales), Hilbert y Bernays consideran los distintos modos como se pueden distribuir los valores veritativos (1 y 0) entre las diversas fórmulas elementales que figuran en cada matriz $\Theta_n$. Sea $q_n$ el número de fórmulas elementales diferentes que figuran en $\Theta_n$. Entonces hay $2^{q_n}$ distribuciones posibles de valores veritativos entre las mismas. Si tomamos dichas fórmulas elementales en el orden de sus *primeras* posiciones (de izquierda a derecha) en $\Theta_n$, cada distribución $\mathscr{W}_n$ queda asociada inequívocamente a un numeral del sistema binario (el sistema de numeración con base 2), a saber, la fila de ceros y unos correspondientes a cada fórmula elemental, en el orden señalado. Llamaré $\#\mathscr{W}_n$ al número representado por ese numeral. Consideremos ahora dos enteros positivos $m$ y $n$, tales que $m < n$. El lector recordará que todas

tarla (Dreben y van Heijenoort, en Gödel, CW, I, 53).[25] En todo caso, Gödel nunca ha cuestionado el Axioma de Selección y en la introducción a la tesis previene expresamente que no se inhibirá de utilizar ningún medio matemático de prueba, una advertencia que sólo resulta pertinente en el punto que

---

las fórmulas elementales de $\Theta_m$ aparecen en $\Theta_n$ *antes* que todas las fórmulas elementales que no figuran en $\Theta_m$. Diremos que la distribución $\mathcal{W}_n$ asociada a $\Theta_n$ contiene a la distribución $\mathcal{W}_m$ asociada a $\Theta_m$ ($\mathcal{W}_m \subset \mathcal{W}_n$) si las fórmulas elementales comunes a ambas matrices reciben, respectivamente, los mismos valores en $\mathcal{W}_m$ y $\mathcal{W}_n$. Nótese que si $\#\mathcal{W}_m$, $\#\mathcal{W}'_m$, $\#\mathcal{W}_n$ y $\#\mathcal{W}'_n$ son los números representativos de cuatro distribuciones asociadas a las matrices indicadas por los subíndices y $\mathcal{W}_m \subset \mathcal{W}_n$ y $\mathcal{W}'_m \subset \mathcal{W}'_n$, entonces $\#\mathcal{W}_m \leq \#\mathcal{W}'_m$ implica que $\#\mathcal{W}_n \leq \#\mathcal{W}'_n$. Diremos con Hilbert y Bernays que una distribución de valores veritativos entre las fórmulas elementales de una matriz es una distribución realizadora (*erfüllende Verteilung*) de la misma, si le confiere el valor 1 ("lo verdadero"). Por hipótesis, hay por lo menos una distribución realizadora de $\Theta_n$, para cada $n > 0$. Si $m < n$, $\Theta_n$ es una conjunción de $\Theta_m$ con otras fórmulas. Por lo tanto, cada distribución $\mathcal{W}_n$ realizadora de $\Theta_n$, contiene una distribución $\mathcal{W}_m$ realizadora de $\Theta_m$. Diremos que $\mathcal{W}_m$ es el *m-componente* de $\mathcal{W}_n$. Nótese que, si $m < k < n$ y $\mathcal{W}_k$ es el $k$-componente de $\mathcal{W}_n$, entonces el $m$-componente de $\mathcal{W}_n$ es *también* el $m$-componente de $\mathcal{W}_k$. "Según esto, resulta (*demnach ergibt sich*) que *entre las distribuciones realizadoras de* $\Theta_m$ *tiene que haber por lo menos una que, para todo* $n > m$, *sea el m-componente de una distribución realizadora de* $\Theta_n$" (Hilbert y Bernays, GG, II, 194; cursiva mía). Entre las distribuciones que tienen esta propiedad, para un dado $m > 0$, hay una cuyo número representativo es menor que el de las demás; llamémosla $\mathcal{W}^*_m$. Hilbert y Bernays prueban que, si $m$ y $n$ son dos enteros positivos cualesquiera tales que $m < n$, entonces $\mathcal{W}^*_m$ es el $m$-componente de $\mathcal{W}^*_n$. Para ello, aprovechan lindamente la representación numérica de las distribuciones: Sea $\mathcal{W}_m$ el $m$-componente de $\mathcal{W}^*_n$ y $\mathcal{W}_n$ la distribución realizadora de $\Theta_n$ tal que $\mathcal{W}^*_m$ es el $m$-componente de $\mathcal{W}_n$; entonces, por la misma definición de $\mathcal{W}^*_m$ y $\mathcal{W}^*_n$ tenemos que $\#\mathcal{W}^*_m \leq \#\mathcal{W}_m$ y $\#\mathcal{W}^*_n \leq \#\mathcal{W}_n$; por otra parte, como hice notar arriba, la aseveración conjunta de las condiciones $\mathcal{W}^*_m \subset \mathcal{W}_n$, $\mathcal{W}_n \subset \mathcal{W}^*_n$ y $\#\mathcal{W}^*_m \leq \#\mathcal{W}_m$ implica que $\#\mathcal{W}_n \leq \#\mathcal{W}^*_n$, de modo que $\#\mathcal{W}_n = \#\mathcal{W}^*_n$; por lo tanto, $\mathcal{W}_n = \mathcal{W}^*_n$ y $\mathcal{W}^*_m$ es el $m$-componente de $\mathcal{W}^*_n$. Este resultado determina inequívocamente una secuencia $\mathcal{W}^*_1 \subset \mathcal{W}^*_2 \subset \dots$, que sirve de base a una realización de $\psi$. Pero su impecable demostración constructiva no puede disimular el carácter no constructivo del paso previo que cité en cursiva.

26 "Por último, debe considerarse que el problema aquí tratado no ha surgido (como, digamos, el problema de la consistencia de las matemáticas) con la disputa sobre los fundamentos [entre brouwerianos y hilbertianos], sino que, aunque nunca se hubiera cuestionado la validez sustantiva de la matemática 'ingenua' (*die inhaltliche Geltung der "naiven" Mathematik*), tendría sentido plantearlo *dentro* de ésta (en contraste, por ejemplo, con el problema de la consistencia). Por esto, una restricción de los medios de prueba no parece más urgente aquí que a propósito de cualquier otro problema matemático" (Gödel, CW, I, 64).

estamos considerando.[26]

Con la secuencia infinita de sistemas de realización de todos los niveles, $S_1 \subset S_2 \subset \ldots$, Gödel tiene lo que necesita para probar que ψ es realizable, si lo son todas las matrices $\Theta_n$ y las respectivas fórmulas verifuncionales $\vartheta_n$. Consideremos el conjunto $S = \{w_1, \ldots, w_h, f_1, \ldots, f_k\}$, caracterizado como sigue: (i) cada $w_i$ es un valor veritativo y cada $f_j$ es una función proposicional definida en el conjunto $\mathbb{N}$ de *todos* los números naturales ($1 \leq i \leq h$; $1 \leq j \leq k$); (ii) $w_i = w_{im}$ para algún $m > 0$ (y, por ende, para todos); (iii) hay un $m > 0$ tal que la restricción de $f_j$ a $Z_m$ es igual a $f_{jm}$ (y, por lo tanto, la restricción de $f_j$ a $Z_n$ es igual a $f_{jn}$ para todo $n > m$). Bajo las condiciones prescritas, es claro que, si asignamos la lista de valores $\langle w_1, \ldots, w_h\rangle$ a la lista $\langle X_1, \ldots, X_h\rangle$ de la variables proposicionales de ψ y entendemos que la lista $\langle F_1, \ldots, F_k\rangle$ de los predicados monádicos o poliádicos de ψ representa la lista de funciones $\langle f_1, \ldots, f_k\rangle$, la fórmula ψ expresa una aseveración verdadera. El sistema *S* ofrece así una realización o —como también se dice— un *modelo* de ψ. Con esto hemos terminado de probar que cualquier $\mathfrak{K}$-fórmula de grado 1 es refutable o realizable. Como se vio antes, esta conclusión implica que cualquier fórmula del **CP1** tiene la misma propiedad. Por su misma naturaleza, la prueba ha establecido, de paso, que si una fórmula del **CP1** es realizable, tiene una realización en el dominio de los números naturales. Esto es una forma del Teorema de Löwenheim-Skolem (*vide* nota 10 y Capítulo 3.2).

No voy a explicar el razonamiento, relativamente fácil, mediante el cual Gödel extiende estos resultados al **CP1=**.[27] En el artículo —mas no en la tesis— prueba además el siguiente teorema:

[27] La argumentación es algo más ágil en el artículo que en la tesis, pero la idea central es en ambos textos la misma: tratar el signo = como un predicado diádico que cumple las condiciones estipuladas en los axiomas **7** y **8** e invocar la suficiencia del **CP1**. Esbozo a continuación el razonamiento del artículo. Sea ϕ una fórmula del **CP1=** que contenga precisamente las variables individuales $\xi_1, \ldots, \xi_h$ y los predicados $\Pi_1, \ldots, \Pi_k$. Construimos la fórmula $\phi' = \phi \wedge \forall x(x = x) \wedge \psi$, donde ψ es una conjunción formada como sigue: si Π es un predicado *n*-ádico ($n \geq 1$) contenido en la lista $\langle \Pi_1, \ldots, \Pi_k\rangle$ y $\langle \zeta_1, \ldots, \zeta_n\rangle$ es cualquier *n*-tuplo formado con variables de la lista $\langle \xi_1, \ldots, \xi_h\rangle$, la fórmula $\forall\xi_r\forall\zeta_1\ldots\forall\zeta_s\ldots\forall\zeta_n(\zeta_s = \xi_r \rightarrow (\Pi\zeta_1\ldots\zeta_s\ldots\zeta_n \rightarrow \Pi\zeta_1\ldots\xi_r\ldots\zeta_n))$ es una de las subfórmulas unidas en ψ por el signo ∧ ($1 \leq r \leq h$; $1 \leq s \leq n$). Si tratamos el signo = como un predicado binario cualquiera, ϕ′ es una fórmula del **CP1** y, por lo tanto, es realizable o refutable. Supongamos que ϕ′ es refutable, o sea, que ¬ϕ′ es deducible de los axiomas **1**–**6**. Entonces, como $\forall x(x = x) \wedge \psi$ es obviamente deducible de los axio-

TEOREMA X. Para que un conjunto infinito numerable de fórmulas del **CP1=** sea realizable, es necesario y suficiente que cada subconjunto finito del mismo sea realizable.

En otras palabras, como se dice hoy, el **CP1=** es *compacto*. Los argumentos utilizados en los primeros pasos de la prueba del Teorema II sirven para mostrar que el Teorema X vale en general, si vale para cualquier conjunto numerable de $\mathfrak{K}$-fórmulas de grado 1. Sea, pues,

$$\Sigma = \{\forall\boldsymbol{\xi}_1\exists\boldsymbol{\zeta}_1\Phi_1(\boldsymbol{\xi}_1;\boldsymbol{\zeta}_1),\forall\boldsymbol{\xi}_2\exists\boldsymbol{\zeta}_2\Phi_2(\boldsymbol{\xi}_2;\boldsymbol{\zeta}_2),\ldots,\forall\boldsymbol{\xi}_n\exists\boldsymbol{\zeta}_n\Phi_n(\boldsymbol{\xi}_n;\boldsymbol{\zeta}_n),\ldots\}$$

un conjunto numerable de $\mathfrak{K}$-fórmulas de grado 1, con $\boldsymbol{\xi}_k$ un $r_k$-tuplo y $\boldsymbol{\zeta}_h$ un $s_h$-tuplo de variables individuales ($k$, $h > 0$). Consideremos ahora, para cada $k > 0$, los $r_k$-tuplos que pueden formarse con las variables individuales $x_0$, $x_1, \ldots$ Ordenados por la suma de sus índices, tales $r_k$-tuplos forman una secuencia $\mathbf{x}_{k1}$, $\mathbf{x}_{k2}, \ldots$ Designaremos con $\mathbf{y}_{hj}$ el $s_h$-tuplo de variables individuales caracterizado por la condición siguiente: al ordenar los $\mathbf{y}_{hj}$ en orden creciente de las sumas $h+j$ (y lexicográficamente, para sumas iguales), la secuencia infinita resultante, $\mathbf{y}_{11}$, $\mathbf{y}_{12}$, $\mathbf{y}_{21}$, $\mathbf{y}_{13}$, $\mathbf{y}_{22}$, $\mathbf{y}_{31}$, $\mathbf{y}_{14}, \ldots$ coincide con la secuencia $x_0$, $x_1, \ldots$ Como se hizo arriba con las matrices $\Theta_n$, definimos una secuencia de matrices $\Xi_1$, $\Xi_2, \ldots,$ mediante las dos estipulaciones siguientes:

$$\Xi_1 = \Phi_1(\mathbf{x}_{11};\mathbf{y}_{11})$$
$$\Xi_n = \Xi_{n-1} \wedge \Xi_1(\mathbf{x}_{1n};\mathbf{y}_{1n}) \wedge \Xi_2(\mathbf{x}_{2(n-1)};\mathbf{y}_{2(n-1)}) \wedge \ldots \wedge \Xi_n(\mathbf{x}_{n1};\mathbf{y}_{n1})$$

---

mas **7** y **8,** $\neg\phi$ es deducible de **1**–**8**. Por otra parte, si $\phi'$ no es refutable, tiene una realización $\mathcal{M}$ en $\mathbb{N}$. La relación aritmética diádica que modela al predicado = en $\mathcal{M}$ tiene que ser reflexiva, simétrica y transitiva (para que satisfaga a $\forall x(x = x) \wedge \psi$); por lo tanto, determina una partición de $\mathbb{N}$ en clases de equivalencia. El conjunto, posiblemente finito, de esas clases de equivalencia provee una realización $\mathcal{M}^*$ de $\phi'$ en la cual cada predicado $\Pi_j$ ($1 \leq j \leq k$) es modelado por la relación que subsiste entre las clases de equivalencia cuyos miembros satisfacen la relación que modela a $\Pi_j$ en $\mathcal{M}$, y el predicado = es modelado por la genuina identidad. Obviamente, la fórmula $\phi$ tiene que estar realizada en toda realización de $\phi'$. En particular, $\mathcal{M}^*$ constituye una realización de $\phi$ entendida como fórmula del **CP1=**.

Sea $(P_n)$ una fila de cuantificadores existenciales que ligan a todas las variables libres en $\Xi_n$. Es claro que $(P_n)\Xi_n$ se deduce del conjunto finito de premisas $\Sigma_n = \{\forall\xi_1\exists\zeta_1\Phi_1(\xi_1;\zeta_1), \forall\xi_2\exists\zeta_2\Phi_2(\xi_2;\zeta_2), \ldots, \forall\xi_n\exists\zeta_n\Phi_n(\xi_n;\zeta_n)\}$ y por consiguiente es realizable si $\Sigma_n$ es realizable. Por lo tanto, si cada subconjunto finito de $\Sigma$ es realizable, también cada matriz $\Xi_n$ es realizable. Por un argumento análogo al aducido en la fase semántica de la prueba del Teorema II se concluye finalmente que si cada matriz $\Xi_n$ es realizable, el conjunto $\Sigma$ es realizable.

Del Teorema X se sigue inmediatamente el siguiente corolario: *Si* $\Sigma$ *es un conjunto numerable de fórmulas del* **CP1=**, *o bien* $\Sigma$ *es realizable*, *o bien* $\Sigma$ *incluye un subconjunto finito de fórmulas* $\{\alpha_1, \ldots, \alpha_n\}$ *tal que* $\alpha_1 \wedge \ldots \wedge \alpha_n$ *es refutable* (Teorema IX; Gödel 1930, p. 357; también en la tesis, § 8, CW, I, 96ss.). Como Gödel expresamente señala, esto significa que, si $\Sigma$ es el conjunto finito o numerablemente infinito de los axiomas de una teoría matemática expresable en el **CP1=**, se cumple una de estas alternativas: o bien (i) la teoría es inconsistente y se puede deducir una contradicción en un número finito de pasos de un número finito de premisas contenidas en $\Sigma$; o bien (ii) hay un modelo —una realización— de la teoría. En una nota, Gödel cita como ejemplo la teoría de los *Grundlagen der Geometrie* de Hilbert, "sin los axiomas de continuidad" (que no se pueden expresar en un cálculo de primer orden).[28] De este modo, el corolario citado resuelve, al menos para este fragmento de la teoría geométrica, la disputa entre Hilbert y Frege a que me referí en el Capítulo 2.1. Hilbert había escrito a Frege, algo livianamente, que "si los axiomas arbitrariamente estipulados, junto con todas sus consecuencias, no se contradicen entre sí, entonces son verdaderos y existen las cosas definidas por ellos: ése es para mí el criterio de la existencia y de la verdad" (Frege, KS, p. 411; *vide* p. 118). En este pasaje, Hilbert posiblemente decía 'existencia' sólo como una manera de hablar, pero la prueba de Gödel le devuelve a la palabra su sentido literal: si la teoría es *sintácticamente*

28 Los axiomas de continuidad forman el grupo V en Hilbert, GG. Son el Postulado de Arquímedes (si *a* y *b* son las distancias entre dos pares de puntos, hay un entero *n* tal que $na > b$) y el llamativo *Axioma de completud* (V,2) agregado por Hilbert en 1902, en virtud del cual ninguna parte propia de un modelo de los axiomas puede a su vez ser un modelo. *Vide supra*, p. 99, nota 27.

consistente, también tiene un modelo que la *realiza*. Irónicamente, el modo no constructivo de razonar empleado por Gödel, aunque cuestionable para un hilbertiano estricto (vgr. Herbrand), no le hubiera merecido reparos a Frege.[29]

El artículo de Gödel contiene otra observación metodológica —ausente en la tesis— que debe haber caído muy bien en la escuela de Hilbert:

> Adviértase que, con vistas al *Entscheidungsproblem*, la equivalencia "válido = deducible" que ha quedado demostrada entraña una reducción de lo no numerable a lo numerable, puesto que "válido" se refiere al conjunto no numerable de las funciones [proposicionales], mientras que "deducible" sólo presupone el conjunto numerable de las pruebas.
>
> (Gödel 1930, p. 356)

Aparentemente, cuando se tomó la molestia de anotar este "avance" hacia la solución positiva del *Entscheidungsproblem*, Gödel no sospechaba que su próximo descubrimiento (1930a, 1931) daría al traste con toda esperanza de

29 En la tesis, Gödel anuncia desde un comienzo este resultado, y se apresura a destacar su interés filosófico. Después de anunciar su propósito en el texto de CW, I, 60 que cité en la nota 3, prosigue: "Se reconocerá fácilmente que esta aseveración equivale a la siguiente: Todo sistema axiomático consistente que conste exclusivamente de [fórmulas del cálculo predicativo de primer orden] tiene una realización (*Realisierung*). (Consistente (*widerspruchslos*) quiere decir aquí que ninguna contradicción puede derivarse mediante un número finito de inferencias formales). La última formulación parece tener cierto interés de por sí, pues la solución de esta cuestión constituye en cierto sentido un complemento teórico del método usual de las pruebas de consistencia (por cierto que sólo para el género especial de sistemas axiomáticos aquí considerado), pues ofrecería una garantía de que este método en todo caso conduce a la meta, es decir, que *o bien tiene que ser posible producir una contradicción, o bien la consistencia tiene que poderse demostrar mediante un modelo*" (Gödel, CW, I, 60).

30 En todo caso, Gödel siguió interesado en el *Entscheidungsproblem* después de 1931. Gödel 1932 prueba que hay un algoritmo para decidir la realizabilidad de cualquier fórmula prenexa de primer orden cuyo prefijo contenga a lo sumo dos cuantificadores universales entre dos filas de cuantificadores existenciales, esto es, cualquier fórmula de la llamada clase $\exists\ldots\exists\forall\forall\exists\ldots\exists$. Este resultado de Gödel mejora el de Ackermann 1928, quien probó que las fórmulas de la clase $\exists\ldots\exists\forall\exists\ldots\exists$ son decidibles. Gödel 1932 demuestra asimismo que el *Entscheidungsproblem* quedaría resuelto positivamente si hubiera un algoritmo para decidir la realizabilidad de la clase $\forall\forall\forall\exists\ldots\exists$ (prefijo con tres cuantificadores universales seguidos de una fila de existenciales). Esto podía verse como un avance enorme hacia la solución del problema general o —según lo entendemos ahora— como una prueba de que la clase $\forall\forall\forall\exists\ldots\exists$ es indecidible.

alcanzarla.[30] Mas, para nosotros que ya lo conocemos, algunas de sus reflexiones parecerían anticiparlo. Algunos pensarán —dice Gödel en la introducción a la tesis— que "la existencia de los conceptos introducidos mediante un sistema de axiomas ha de definirse justamente por su consistencia (*Widerspruchslosigkeit*)."[31]

> Pero esta definición —con sólo que uno exija, como es obvio, que el concepto de existencia así introducido obedezca a las mismas reglas de operación que el elemental— *presupone evidentemente el axioma de la solubilidad de todo problema matemático*. Pues si se demostrara la insolubilidad de algún problema (relativo, digamos, al dominio de los números naturales), de esto se desprendería conforme a dicha definición que *existen* dos realizaciones no isomórficas del sistema axiomático de los números reales, mientras que, por otra parte, se puede demostrar el isomorfismo que cualquier par de realizaciones. Pero la demostración de la insolubilidad de un problema no se puede excluir de antemano, si se piensa que hablamos *sólo de insolubilidad con ciertos modos de inferencia formal precisamente señalados*. Pues todos los conceptos que vienen a cuento aquí (demostrable, consistente, etc.) sólo tienen un sentido exacto si se deslindan con precisión los modos de inferencia permitidos.
>
> (Gödel, CW, I, 60–62; cursiva mía)[32]

El artículo aborda brevemente la cuestión de la independencia de los axiomas **1**–**8**. Recuerda que Bernays (1926) ya había probado que ninguno de los axiomas **1**–**4** puede deducirse de los otros tres. Obviamente, su independencia no se altera al agregar los axiomas **5**–**8**. Para mostrar que también éstos son independientes —en el sistema total— Gödel opera sustituciones o adopta postulados que, en cada caso, invalidan a uno de ellos sin afectar la

31 De paso señalo que el pronombre posesivo 'ihre' ('de ellos' o 'de ella') que precede a 'Widerspruchslosigkeit' ('ausencia de contradicción', arriba traducido 'consistencia') sólo puede referirse al sustantivo plural 'Begriffe' ('conceptos') o al femenino singular 'Existenz' ('existencia'), y no al neutro singular 'Axiomensystem' ('sistema de axiomas') como cree el traductor inglés (Gödel, CW, I, 61). Por lo demás, en el texto alemán 'Begriffe' está más cerca de 'ihre' que los otros dos sustantivos. Habría que explicar, claro está, qué significa la 'ausencia de contradicción' de un grupo de conceptos. Por esto, he preferido dar una traducción ambigua.

32 En el original sólo van en cursiva las palabras correspondientes a 'formal precisamente señalados'.

validez de los restantes (y tampoco, por ende, la de las fórmulas deducibles de éstos). Por ejemplo, para mostrar que **8** no se deduce de **1**–**7** postula que $\forall x \forall y(x = y)$, lo que hace falso a **8** en cualquier dominio con dos o más objetos, pero es perfectamente compatible con **1**–**7**. En el pasaje paralelo de la tesis Gödel examina asimismo la indispensabilidad de las cuatro reglas de inferencia, aunque algo apresuradamente, me parece (de *modus ponens* y la regla de sustitución de variables no dice más que esto: "que las reglas 1 y 2 no son superfluas es algo perfectamente trivial (*ist ja vollkommen trivial*)"— CW, I, 96). En ninguno de estos textos se sugiere ni por un instante que la independencia de los axiomas de la lógica abra la posibilidad de contemplar lógicas alternativas. Sólo se trata de mostrar que los principios cuya *suficiencia* para deducir todas las verdades lógicas enunciables en el **CP1**= ha quedado establecida son también *imprescindibles* para ello. Al fin y al cabo, si la mera deducibilidad en un sistema formal arbitrario fuese un criterio de verdad lógica ("en el sistema") el problema resuelto por Gödel en la parte principal de estos trabajos no tendría sentido.

## 2.9 EL PROGRAMA DE HILBERT VISTO MÁS DE CERCA

Se dice comúnmente que el segundo gran descubrimiento de Gödel —sus famosos teoremas de "incompletud" (1930a, 1931)— significa el fracaso del programa de Hilbert. Estudiaremos el hallazgo de Gödel y sus consecuencias más directas en los Capitulos 2.10 y 2.11. Pero para juzgar su efecto sobre el programa de Hilbert conviene primero examinar los términos en que él mismo lo explicaba y promovía hasta fines de 1930, antes del *shock* causado por Gödel.

### *2.9.1 Axiomatización y formalización*

En los planes de Hilbert para fundamentar sólidamente las matemáticas hay dos aspectos que suelen confundirse: la *axiomatización* y la *formalización*. Mediante la axiomatización de una teoría matemática —como la practican ejemplarmente Pasch (1882) y Hilbert (1899, 1900)— se trata de caracterizar exhaustiva y definitivamente el objeto de la misma mediante una lista de enunciados —los axiomas y definiciones— de suerte que *todo* lo que la teoría tenga que decir acerca de dicho objeto se pueda demostrar *razonando* a partir de esa lista. Mediante la formalización —según la concibe y ejecuta Frege (1879, 1893, 1903)— se trata de representar las aseveraciones de la teoría en una "escritura conceptual" de sintaxis sencilla y rigurosa, diseñada para facilitar la presentación de los razonamientos demostrativos en una forma canónica, fácilmente controlable.

Esta descripción sumaria nos da criterios para medir los aciertos y los límites de una axiomatización o una formalización. Consideremos una teoría matemática tradicional cualquiera, por ejemplo, la geometría euclidiana ordinaria que Hilbert axiomatizó en 1899. La llamaré $\mathscr{E}$. *Todo* lo que $\mathscr{E}$ tiene que decir sobre su objeto, es decir, todos los *teoremas* de $\mathscr{E}$ se pueden, por

cierto, demostrar razonando a partir del solo axioma siguiente: 'cero no es idéntico a cero' (sea lo que fuere el objeto aquí denominado 'cero'). Pero este axioma implica también la *negación* de cada uno de esos teoremas. Para que *caracterice* a $\mathscr{E}$, es necesario que la lista de axiomas propuesta sea *consistente*: si implica una cierta aseveración *p*, no debe implicar la negación de *p*. Por otra parte, la consistencia de la lista de axiomas no es suficiente para caracterizar a $\mathscr{E}$. Es preciso además que ella sea *completa*, esto es, que implique *todos* los teoremas de $\mathscr{E}$. Hilbert menciona estos dos requisitos de una buena axiomatización a propósito, justamente, de su propia axiomatización de $\mathscr{E}$ (1900, en GG[7], p. 242). El requisito de *consistencia* es crítico e insoslayable. Pero el requisito de *completud* se puede trivializar —en el caso de una teoría nueva o de una tradicional cuyo alcance se disputa— *decretando* que el objeto de la teoría que se pretende axiomatizar es *precisamente* aquello que los axiomas caracterizan. Es lo que hace Hilbert a propósito de los números reales. En la segunda mitad del siglo XIX éstos habían sido definidos por Cantor y otros autores de varias maneras que apelaban a la existencia de colecciones infinitas bien determinadas de números racionales (cf. Capítulo 1.3). Kronecker negaba que hubiera tales colecciones. Hilbert (1900) coge el toro por las astas: formula una lista de cuatro grupos de axiomas que caracterizan lo que hoy se llama un *cuerpo ordenado, completo y arquimédico* (el lector que no conozca este concepto debe ahora mismo consultar el Glosario, *s.v.* 'cuerpo') y propone la estipulación siguiente:

> En adelante no debemos entender por 'el conjunto de los números reales' algo así como la totalidad de todas las leyes posibles conforme a las cuales se suceden los elementos de una secuencia fundamental [definida en la p. 16—R.T.], sino más bien —como acaba de explicarse— un sistema de cosas cuyas relaciones mutuas están dadas por el sistema *finito y cerrado* de los Axiomas I–IV, y sobre las cuales valen nuevas aseveraciones solamente si se las puede derivar de esos axiomas mediante un número finito de inferencias lógicas.
>
> (Hilbert 1900, en GG[7], p. 246)

La frase de Hilbert "derivar de esos axiomas mediante un número finito de inferencias lógicas" ("mittels einer endlichen Anzahl von logischen Schlüssen aus jenen Axiomen ableiten") corresponde a la que usé arriba: "demostrar razonando a partir de [ellos]". Años más tarde, Hilbert ridiculizará a quienes

"subrayan la exigencia de que en la matemática rigurosa se permite en una demostración sólo un número *finito* de inferencias—como si ya alguien hubiese logrado ejecutar una vez infinitas inferencias" (1925, p. 162). Esta atinada observación no se hace cargo, empero, de la vaguedad inherente a la idea misma de un *número de inferencias*. Pues ¿qué es *una* inferencia? Todos hemos pasado por la experiencia de oir o leer a un profesor de matemáticas que declara "inmediato" el paso de unas premisas a una conclusión que no nos parece nada de evidente. Si luego, con trabajo, conseguimos entenderlo, será porque logramos analizarlo en *varias* inferencias que sí son simples *para nosotros*.

La formalización de las matemáticas elimina la incertidumbre de tales análisis. En la matemática formalizada, cada razonamiento correcto se representa —como hemos visto— mediante una listas de fórmula que se suceden unas a otras conforme a ciertas reglas de cómputo, y cada aplicación simple de una de estas reglas cuenta como *una* inferencia. Pero Hilbert ve en la formalización, sobre todo, un camino para asegurar el cumplimiento del requisito de consistencia. Él había probado (en GG, § 9) que su teoría axiomática de la geometría era *por lo menos tan consistente como* la aritmética (incluida la teoría de los números reales). Pero en el caso de la aritmética misma evidentemente no cabe apelar a la consistencia de otra teoría más fundamental.

### *2.9.2 Balbuceos formales (Hilbert 1904)*

Hilbert esboza —muy imperfectamente— una formalización de la aritmética en su conferencia "Sobre los fundamentos de la lógica y la aritmética" (1904). Empieza explicando las dificultades de una fundamentación de la aritmética a través de un breve comentario crítico de las opiniones de seis autores: Kronecker, Helmholtz, Christoffel, Frege, Dedekind y Cantor. Hilbert concede al primero el mérito de haber visto "en el concepto del número entero el fundamento genuino de la aritmética". Pero Kronecker pensaba que el número entero nos estaba dado directa e inmediatamente *como concepto general*, lo cual le impidió reconocer que ese concepto necesita y admite una

fundamentación.[1] Helmholtz sustentaba una filosofía empirista de la aritmética, que según Hilbert queda refutada por cuanto "nunca se podrá colegir de la experiencia —vale decir, mediante un experimento— la posibilidad o la existencia de un número arbitrariamente grande", dado que "el número de las cosas que son objeto de nuestra experiencia, por grande que sea, es menor que cierta cota finita" (GG[7], p. 248). Hilbert nombra a Christoffel como el principal de los "oportunistas" que entendieron correctamente que los irracionales eran indispensables para el análisis y defendieron su existencia contra Kronecker, sin lograr refutarlo.

Para Hilbert, los tres autores restantes son quienes "han calado más hondo en la esencia del número entero". Frege, que se propuso "fundar las leyes de la aritmética por medio de la lógica, concebida en el sentido tradicional", tiene el mérito de "haber reconocido correctamente las propiedades esenciales del concepto de número entero así como el significado de la inferencia por inducción completa". Pero su intento naufraga en paradojas debido a que Frege (i) da por supuesto que "un concepto (un conjunto) se ha definido y se puede utilizar inmediatamente en cuanto se haya determinado para cada objeto si cae o no bajo el concepto" y (ii) "no somete el concepto 'todo' (*'jeder'*) a restricción alguna". El fracaso de Frege muestra que "las concepciones y medios de investigación de la lógica, concebida en sentido tradicional, no están a la altura de las rigurosas exigencias de la teoría de conjuntos". Hilbert proclama enfáticamente que "*las investigaciones sobre el concepto de número tienen que proponerse de antemano como una meta principal evitar tales contradicciones y aclarar esas paradojas*".[2] En la obra de

1 Más tarde, Hilbert describirá a Kronecker como su precursor más directo: éste "formuló claramente y dilucidó con numerosos ejemplos una concepción que hoy coincide en lo esencial con nuestra postura finita" (Hilbert 1931, p. 487). Como bien señala Bernays, el acuerdo que Hilbert percibe entre su propio modo de pensar y el de Kronecker estriba "especialmente en el concepto intuitivo de numeral y en el reconocimiento de que la forma intuitiva de la inducción completa, esto es, el modo de inferencia que se basa en la representación intuitiva de la 'construcción' de los numerales, es evidente y no requiere ulterior fundamentación" (1935, p. 203). Sobre la diferencia entre dicha "forma intuitiva de la inducción completa" y la inducción matemática irrestricta, véase más adelante la nota 15.

2 El párrafo dedicado a Frege figura en Hilbert, GG[7], pp. 248s. Sorprende que Hilbert creyera que Frege concebía la lógica "en sentido tradicional" (*in hergebrachtem Sinne*). Merece atención la sugerencia de que las paradojas vienen de la aceptación indebida

Dedekind, Hilbert destaca el método "trascendental" aplicado a la demostración de la existencia del infinito (bosquejada arriba, en la p. 154, nota 12). Porque esa demostración recurre al concepto —contradictorio— de la totalidad de las cosas, Hilbert se niega a aceptarla como "viable y segura". Por último, Cantor sintió la contradicción y expresó, según Hilbert, ese sentimiento al distinguir entre "conjuntos (*Mengen*) 'consistentes' y 'no consistentes'"; pero no basó este distingo en ningún criterio preciso, con lo cual "deja libertad al arbitrio subjetivo y por ende no provee una seguridad objetiva".[3]

Completado este panorama, Hilbert declara estar convencido de que "las dificultades mencionadas se dejan superar y que es posible lograr una fundamentación rigurosa y plenamente satisfactoria del concepto de número" mediante un método que él llama *axiomático*. A la luz del bosquejo presentado en el resto de la conferencia, es claro que se trata de un intento, todavía balbuceante, de *formalización* de la aritmética de los enteros positivos. No es posible reproducirla aquí en todos sus detalles,[4] pero vale la pena llamar la atención sobre algunos de ellos. La exposición de Hilbert comienza con la siguiente estipulación:

> Llámese *cosa pensada* (*Gedankending*) o, brevemente, *cosa* (*Ding*) a un objeto de nuestro pensamiento y nómbresela mediante un signo.
> (Hilbert 1904, en GG[7], p. 250)

Hilbert propone la consideración de cinco cosas, nombradas mediante los signos 1, =, $\mathfrak{u}$, $\mathfrak{f}$, y $\mathfrak{f}'$. Estas cosas —que luego (p. 255) llama "simples"— se *combinan* en lo que, si damos por supuesto el concepto de número natural finito, describiríamos sin vacilar como *n*-tuplos o listas. Pero Hilbert quiere

de ciertos *conceptos*, no de la suposición injustificada de que *cualquier* concepto constituye un *conjunto*, esto es, un *objeto* que a su vez es concebible y por ende —bajo dicha suposición— es *elemento* de algún conjunto.

3 Hilbert 1904, en GG[7], pp. 249. Como vimos en el Capítulo 1.6, Cantor hablaba de *pluralidades* (*Mannigfaltigkeiten*) consistentes e inconsistentes. Sólo las *consistentes* son *conjuntos* (*Mengen*). Es curioso que Hilbert desdeñara cuando hacía historia la exactitud que juzgaba indispensable en matemáticas.

4 Hay una traducción inglesa de Hilbert 1904 en van Heijenoort 1967, pp. 129–138.

tal vez evitar ese supuesto[5] y se limita por eso a ilustrar tales combinaciones de cosas mediante concatenaciones de los signos que las nombran, sin la más mínima indicación de cómo hay que entenderlas. He aquí sus ejemplos:

$$1 =, \quad 11 =, \ldots, (1)(= 1)(= = =), ((11)\ (1)\ (=))\ (= =),$$

$$1 = 1,\ (11) = (1)(1)$$

La coma y los puntos suspensivos se usan para separar distintos ejemplos, y el espacio en blanco que a veces precede o sigue al signo = sirve, me imagino, sólo para facilitar la lectura. Pero los paréntesis ) y ( obviamente forman parte de la combinación de signos que nombra una dada combinación de cosas, aunque Hilbert no dice como se usan, ni siquiera los introduce formalmente. Hilbert nos invita enseguida a "pensar que las combinaciones de esas dos cosas simples se reparten en dos clases, la *clase de los existentes* (*der Seienden*) y la de los *no-existentes* (*der Nichtseienden*)" (GG[7], p. 251). Si $a$ es una concatenación de signos que nombra una combinación, $a$ significa también el enunciado de que la combinación nombrada por $a$ pertenece a la clase de los existentes y $\overline{a}$ significa el enunciado de que dicha combinación pertenece a la clase de los no-existentes. El enunciado $a$ es *correcto* (*richtig*) si la combinación $a$ efectivamente pertenece a la clase de los existentes, pero si ella pertenece a la clase de los no-existentes es correcto el enunciado $\overline{a}$. Los enunciados $a$ y $\overline{a}$ constituyen —dice Hilbert— una *contradicción*.

Hilbert enseña luego a formar un enunciado nuevo a partir de dos enunciados dados intercalando entre éstos el signo o. y el signo u. (que funcionan, respectivamente, como el signo de disyunción $\vee$ y el signo de conjunción $\wedge$). Si $A$ y $B$ son enunciados, $A|B$ se lee 'de $A$ se sigue $B$' o 'si $A$ es correcto, también es correcto $B$'. Esta explicación y el empleo ulterior del signo | en la formulación de reglas de inferencia (GG[7], p. 260) sugiere que dicho signo corresponde a nuestro símbolo (metalingüístico) de deducibilidad $\vdash$. Pero Hilbert trata a $A|B$ como enunciado formal, en que el signo | corresponde a nuestra flecha $\rightarrow$. Curiosa es también la forma como Hilbert intro-

[5] Más adelante Hilbert reconoce, sin embargo, que para demostrar la consistencia del sistema formal esbozado necesitará el concepto de número ordinal finito (GG[7], p. 255).

duce lo que hoy llamaríamos cuantificadores. Si $A(x)$ es un enunciado en que uno o más símbolos que designan cosas se han reemplazado por la variable $x$ y $A_1, A_2, A_3, \ldots$ es una lista de todos los enunciados diferentes que pueden formarse reemplazando uniformemente $x$ en $A(x)$ por una concatenación de signos 1 y =, entonces $A(x^{(o)})$ representa la disyunción infinita $A_1 \vee A_2 \vee A_3 \vee \ldots$ y se lee 'por lo menos para un $x$, $A$' y $A(x^{(u)})$ representa la conjunción infinita $A_1 \wedge A_2 \wedge A_3 \wedge \ldots$ y se lee 'para cada $x$, $A$'.[6]

A continuación, Hilbert postula cinco "axiomas". Son concatenaciones de los cinco signos 1, =, $\mathfrak{u}$, $\mathfrak{f}$, y $\mathfrak{f}'$ y una o más de las letras $x$, $y$, $w$ ($w$ figura sólo en las combinaciones $w(x)$ y $w(y)$). El uso de estas letras se explica así: $x$ e $y$ representan a uno de los signos 1 y =, o una combinación cualquiera de estos signos; $w(x)$ representa una combinación arbitraria de los cinco signos que contiene la combinación de los signos 1 y = representada por la letra $x$. He aquí los axiomas, tal como Hilbert los escribe:

1 $\quad x = x$

2 $\quad \{x = y \text{ u. } w(x)\} \mid w(y)$

3 $\quad \mathfrak{f}(\mathfrak{u}x) = \mathfrak{u}(\mathfrak{f}'x)$

4 $\quad \mathfrak{f}(\mathfrak{u}x) = \mathfrak{f}(\mathfrak{u}y) \mid \mathfrak{u}x = \mathfrak{u}y$

5 $\quad \overline{\mathfrak{f}(\mathfrak{u}x) = \mathfrak{u}1}$

6 El original es menos claro que mi paráfrasis. Lo cito a continuación, en aras de la exactitud histórica. Adviértase que en las páginas que preceden a este pasaje Hilbert no ha mencionado en ningún momento la posibilidad de que un enunciado (*Aussage*) contenga una variable (*Willkürliche*):

> Wir wollen unter $A_1, A_2, A_3, \ldots$ diejenigen Aussagen verstehen, die —kurz ausgedrückt— aus einer Aussage $A(x)$ hervorgehen, indem wir an Stelle der „*Willkürlichen*" $x$ die Gedankendinge 1, =, und die Kombinationen derselben nehmen; dann schreiben wir die Aussagen
>
> $$A_1 \text{ o. } A_2 \text{ o. } A_3, \ldots \quad \text{bzw.} \quad A_1 \text{ u. } A_2 \text{ u. } A_3, \ldots$$
>
> auch wie folgt:
>
> $A(x^{(o)})$, in Worten: „wenigstens für $x$"
> bzw. $A(x^{(u)})$, in Worten: „für jedes einzelene $x$";
>
> hierin erblicken wir lediglich eine abkürzende Schreibweise.
>
> (Hilbert 1904, en GG[7], p. 252)

Se trata, evidentemente, de lo que hoy llamamos esquemas. Los axiomas propiamente tales serán las infinitas concatenaciones de signos que pueden formarse reemplazando *x*, *y*, *w*(*x*) y *w*(*y*) según las indicaciones ofrecidas. Por eso desconcierta leer que **1** y **2** son “enunciados (*Aussagen*)” que conjuntamente “constituyen la *definición* del *concepto* = (igual)” ($GG^7$, p. 252). Las explicaciones dadas por Hilbert acerca de los tres restantes sugieren que ellos deben hacer las veces de los axiomas **P1**–**P4** de Peano (Sección 2.3.1). Los objetos $\mathfrak{u}$, $\mathfrak{f}$, y $\mathfrak{f}'$ se llaman, respectivamente “conjunto infinito (*unendliche Menge*)”, “siguiente (*Folgendes*)” y “operación concomitante (*begleitende Operation*)”. La combinación $\mathfrak{u}x$ “se llama” un *elemento* del conjunto infinito $\mathfrak{u}$.

> Entonces, el axioma **3** expresa que a cada elemento $\mathfrak{u}x$ le sigue una determinada cosa pensada $\mathfrak{f}(\mathfrak{u}x)$, que es igual a un elemento del conjunto $\mathfrak{u}$, a saber, al elemento $\mathfrak{u}(\mathfrak{f}'x)$, que también pertenece al conjunto $\mathfrak{u}$. El axioma **4** expresa el hecho de que, si los elementos que siguen a dos elementos del conjunto $\mathfrak{u}$ son iguales, esos dos elementos también son iguales. Conforme al axioma **5** no hay en $\mathfrak{u}$ ningún elemento al que siga el elemento $\mathfrak{u}1$; por eso este elemento $\mathfrak{u}1$ se llama el primer elemento de $\mathfrak{u}$.
> (Hilbert 1904, en $GG^7$, p. 254)

Es difícil imaginarse cómo Hilbert esperaba fundamentar la aritmética con estos cinco (tipos de) axiomas y las dos reglas de inferencia que formula explícitamente.[7] Pero la idea central de su programa para una prueba directa de consistencia ya se insinúa claramente. Subraya que el axioma **5** es el único con la forma $\overline{a}$, propia de un enunciado que adjudica una combinación de objetos simples a la clase de los no existentes. El sistema sería inconsistente si permitiera inferir un enunciado que adjudicara a la clase de los existentes una de las combinaciones así clasificadas por **5**; esto es, si permitiera inferir el enunciado $\mathfrak{f}(\mathfrak{u}x^{(o)}) = \mathfrak{u}1$.[8] Ahora bien, este enunciado ciertamente *no es* lo

7 Para mayor claridad, doy las dos reglas en nuestro simbolismo: (i) de $((a \to b) \wedge (\neg a \to b))$ inferir $b$; (ii) de $((a \vee b) \wedge (a \vee c))$ inferir $(a \vee (b \wedge c)$. Adviértase que ninguna de las dos envuelve expresiones con cuantificadores. Tampoco hay una regla de sustitución, pero esto es natural si todos los axiomas se dan en la forma de esquemas.

8 Hilbert dice que “los enunciados que forman con **5** una contradicción tienen que ser [ . . . ] de la forma: **6** $\mathfrak{f}(\mathfrak{u}x^{(o)}) = \mathfrak{u}1$” ($GG^7$, p. 254). Si el cuantificador existencial con-

que Hilbert llama una ecuación homogénea, esto es, una combinación de la forma $a = b$, en que $a$ y $b$ son $n$-tuplos de objetos simples (para un mismo $n$). En virtud de esto, para completar la prueba de consistencia basta demostrar que de los axiomas **1**–**4** no puede inferirse ninguna ecuación que no sea homogénea. Para efectuar dicha demostración hacen falta, según Hilbert, el concepto de número ordinal finito y ciertos teoremas sobre el concepto de equinumerosidad que —nos dice— pueden formularse y derivarse sin esfuerzo (*ohne Mühe*). Seguramente tiene razón. En los años 20 —como ya hemos visto— varios discípulos de Hilbert probarán rigurosamente y sin dificultad la consistencia del muñón de aritmética basado en los axiomas **P1**–**P4** de Peano.[9]

---

tenido en **6** se concibe a nuestro modo, es claro que (a diferencia de **5**) **6** no es una mera "forma" de enunciado (un esquema), sino un enunciado genuino. Pero no podemos saber con precisión cómo lo concebía Hilbert en 1904.

9 Hilbert 1904 concluye con ciertos "principios normativos de la construcción y ulterior desarrollo de las leyes del pensamiento matemático en la forma proyectada", numerados del I al V. Se trata más bien de consideraciones metodológicas. Doy enseguida una paráfrasis abreviada. I. Cualquier enunciado debe reconocerse como correcto en cuanto se establezca que no resultará una contradicción al añadirlo como axioma a los enunciados probadamente correctos. (Según Hilbert esto expresa el principio creador que nos autoriza a formar libremente conceptos nuevos sin otra restricción que la de evitar las contradicciones). II. En los axiomas, el recorrido de las variables de cuantificación se limita a las cosas que se dieron inicialmente por sentadas y sus combinaciones, así como a las cosas que luego se definan en términos de aquéllas. (Se recordará que en el malogrado sistema de Frege, cada variable de cuantificación recorría supuestamente la totalidad absoluta de las cosas). III. Al revés de la concepción corriente (a principios de siglo), el concepto de *elemento de un conjunto* depende del concepto de *conjunto*: "El *conjunto* se define en general como una cosa pensada $m$ y las combinaciones $mx$ se llaman elementos del conjunto $m$" (GG[7], p. 258). IV. Al investigar un dado sistema de axiomas con arreglo a estos principios hay que repartir en dos clases las combinaciones de las cosas que se dieron por sentadas, la clase de los existentes y la de los no-existentes, tratando a los axiomas como prescripciones que esta clasificación tiene que satisfacer. V. Cuando se habla *en plural* de cosas pensadas y sus combinaciones, de variables, de clases de combinaciones, uno se refiere siempre a un número limitado de tales objetos.

### *2.9.3 Teoría de la prueba*

> *»Wir kennen die Unendlichkeit aus der Beschreibung.« Nun, dann gibt es eben diese Beschreibung und nichts sonst.*
>
> *»Conocemos lo infinito por la descripción.« Bien, entonces lo que hay es justamente esa descripción y nada más.*
>
> WITTGENSTEIN, PB § 136

Hilbert nunca elaboró en detalle la formalización de la aritmética esbozada en su conferencia de 1904.[10] Su próxima publicación sobre el tema —"Pensamiento axiomático" (1918)— remite a la obra de Russell, que propone completar con una prueba de consistencia. En este escrito —una conferencia pronunciada en Zurich en septiembre de 1917— 'método axiomático' no es un sinónimo de 'formalización' (como en Hilbert 1904), sino, como es justo, de lo que he llamado 'axiomatización'. Hilbert señala que los hechos comprendidos en el campo de estudio de cualquier especialidad científica se ordenan con ayuda de "un cierto *andamio de conceptos* (*Fachwerk von Begriffen*) de tal modo que al objeto particular del campo de estudios le corresponde un concepto de ese andamiaje y a cada hecho dentro del campo de estudios le corresponde una relación lógica entre los conceptos" (Hilbert 1918, en GA III, 146). Dicho andamio conceptual es la *teoría* de ese campo. Su construcción descansa en "unas pocas aseveraciones destacadas, [que] bastan por sí solas para construir a partir de ellas todo el andamio, conforme a principios lógicos" (GA III, 147). Tales aseveraciones son los *axiomas* de la teoría. Para que la teoría sirva de veras para ordenar los hechos y orientar el pensamiento es menester, según Hilbert, hacerse una idea acerca de la mutua dependencia e independencia de sus aseveraciones y tener una garantía

10 Bernays 1935, p. 200, n. 1, observa que el matemático húngaro Julius König, en su inconclusa *Lógica sintética* —publicada póstumamente como *Nuevos fundamentos de la lógica, aritmética y teoría de conjuntos* (König 1914)— continúa la línea de investigación iniciada en Hilbert 1904. Bernays agrega que con esa obra König fue un precursor de la teoría de la prueba, pero que ella no ejerció ninguna influencia sobre Hilbert. En cambio, el estudio del húngaro von Neumann "Sobre la teoría hilbertiana de la prueba" (1927) sí se vincula a la obra de König.

de su consistencia (*Widerspruchslosigkeit*). Hilbert cita varios ejemplos de dependencia e independencia de asertos en teorías conocidas[11] y aborda luego la cuestión de la consistencia, cuya "suprema importancia" es patente, pues "la presencia de una contradicción en una teoría obviamente pone en peligro la subsistencia de toda la teoría" (GA III, 150). Hilbert alude una vez más a las pruebas de consistencia relativa —como la que demuestra que su teoría axiomática de la geometría no puede contener contradicciones si la teoría de los números reales es consistente— y recuerda que Weierstraß y Dedekind, empleando conceptos conjuntistas, lograron reducir la consistencia de la teoría de los reales a la consistencia de la teoría de los enteros. Pero en el caso de la teoría de los enteros y de la teoría de conjuntos ese procedimiento de reducción a una teoría diferente es inviable "porque fuera de la lógica no hay ninguna otra disciplina a la que uno pueda entonces apelar" (GA III, 153). Como "el examen de la consistencia es una tarea insoslayable" habrá que "axiomatizar la lógica misma y demostrar que la teoría de los números y la teoría de conjuntos sólo son partes de la lógica". Esta tarea, preparada por "las profundas investigaciones de Frege", ha sido acometida con muchísimo éxito por "el agudo matemático y lógico Russell".

> En la consumación de esta grandiosa empresa russelliana de la *axiomatización de la lógica* podría verse la coronación de la obra de la axiomatización en general. Pero esa consumación demandará aún nuevas y variadas labores. Una consideración más atenta nos revela enseguida que en el caso de los números enteros y los conjuntos la cuestión de la consistencia no

[11] Hilbert destaca especialmente la independencia del Postulado de Arquímedes en la física: "El hecho de que añadiendo distancias terrestres alcancemos las dimensiones y distancias entre los cuerpos en los espacios siderales [ . . . ], así como el hecho de que las distancias en el interior del átomo se puedan expresar en metros, no son de ningún modo una mera consecuencia lógica de los teoremas sobre congruencia de triángulos y la configuración geométrica, sino un resultado de la investigación empírica" (GA III, 149). La siguiente versión del Postulado expresa bien, según él, la forma como los físicos lo han utilizado siempre, aunque sin mencionarlo:

> Si se le asigna un cierto grado arbitrario de precisión a la validez de un enunciado físico, es posible indicar pequeños dominios (*Bereiche*) dentro de los cuales las condiciones del enunciado pueden variar libremente sin que la discrepancia resultante exceda la imprecisión admisible.
>
> (Hilbert 1918, en GA III, 150]

> está aislada sino que pertenece a un vasto repertorio de dificilísimas cuestiones epistemológicas de específica coloración matemática. Menciono, para caracterizar brevemente este campo de problemas, la cuestión de la *solubilidad* en principio *de todo problema matemático*, el problema de la posibilidad ulterior de *controlar* el resultado de una investigación matemática, también la pregunta por un *criterio de simplicidad* de las demostraciones matemáticas, la cuestión de la relación entre *sustantividad* (*Inhaltlichkeit*) *y formalismo*, y por último el problema de la *decidibilidad* de una cuestión matemática mediante un número finito de operaciones.
>
> (Hilbert 1918, en GA III, 153)

Los problemas que Hilbert enumera se han comentado en los capítulos anteriores. Forman parte, según él, del tema de una disciplina matemática especial. Para dominarlo "debemos investigar el concepto mismo de la prueba específicamente matemática (*den Begriff des spezifischen mathematischen Beweises*): así como el astrónomo tiene en cuenta el movimiento de su observatorio, el físico debe prestar atención a la teoría de su aparato y el filósofo critica a la razón misma" (GA III, 155). Esta es la disciplina que más tarde llamará *teoría de la prueba* (*Beweistheorie*: Hilbert 1928 en GG[7], p. 289) y también *metamatemática* (*Metamathematik*: Hilbert 1923 en GA, III, 179). A ella pertenecen las investigaciones de Ackermann (1924), von Neumann (1927) y Herbrand (1929, 1930, 1931) estudiadas en los Capítulos 2.7 y 2.8. Por esos mismos años, Hilbert difunde las ideas básicas de la teoría de la prueba y defiende su valor para la fundamentación de las matemáticas, a través de la docencia y en una serie de alocuciones en seminarios y congresos (1922, 1923, 1926, 1928, 1928a, 1931). En el resto de esta sección presentaré esas ideas a la luz de estos textos, prescindiendo de los detalles formales, que creo haber ilustrado suficientemente en dichos capítulos.

Hilbert describe varias veces el "supuesto mínimo"[12] indispensable para poner en marcha el discurso lógico-matemático (así como el pensamiento científico en general):

> Algo nos está ya dado de antemano en la representación; ciertos objetos concretos extralógicos que preceden como vivencia inmediata a todo pensamiento. Para que la inferencia lógica sea segura, estos objetos tienen

12 "Das geringste Maß von Voraussetzung (la mínima medida de presuposición)"—Hilbert 1928, en GG[7], p. 290.

> que dejarse abarcar con la mirada (*überblicken*) en todas sus partes, y su presentación (*Aufweisung*), su distinción, su sucesión o concatenación (*ihr Aufeinanderfolgen oder Nebeneinandergereihtsein*) está dado directa e intuitivamente junto con los objetos como algo que no se deja reducir a otra cosa ni requiere una reducción.
>
> (Hilbert 1931, p. 486)[13]

A primera vista, parecería que se trata de un supuesto empírico. Es un hecho común que percibimos objetos que podemos distinguir, reconocer y combinar; sin ir más lejos: las letras impresas en esta página. A tono con ello, Hilbert propone —"en directa oposición a Frege y Dedekind"— una aritmética cuyos objetos sean "los mismos signos, cuya figura podemos reconocer generalmente y con seguridad independientemente del lugar y el tiempo y de las condiciones particulares de la producción del signo, así como de diferencias insignificantes en su ejecución" (1922, en GA III, 163). Pero al considerar cómo Hilbert usa su "supuesto mínimo" vemos que éste rebasa los límites de cualquier experiencia. A continuación del último pasaje citado, da la siguiente definición de 'número' (*Zahl*):

> El signo 1 es un número.
> Un signo que empieza con 1 y termina con 1 y en el cual cada 1 va seguido por [el signo] + y cada + va seguido por 1 también es un número; vgr. los signos
>
> 1 + 1
> 1 + 1 + 1
>
> (Hilbert 1922, en GA III, 163)

Fuera de estos "signos" sin significado, que constituyen el objeto propio de la aritmética, Hilbert introduce signos con significado, como = y >, que sirven para trasmitir aseveraciones (además introduce abreviaturas —2, 3, . . . — que significan números). Una fórmula que consta de dos números separados por el signo = expresa una aseveración verdadera si y sólo si el número a la derecha de = reproduce el número a la izquierda. Una fórmula que consta de dos números separados por el signo > expresa una aseveración verdadera si y sólo si el número a la izquierda de > reproduce el número a la derecha

[13] Cf. 1922, en GA III, 162s.; 1926, p. 170; 1928 en GG[7], pp. 289s.

seguido de una o más copias del signo '+1'. Tras esta concisa explicación, Hilbert propone usar letras góticas como numerales (*Zahlzeichen*), esto es como variables que representan números indeterminados. Demuestra luego el teorema siguiente: *cualesquiera que sean los números* $\mathfrak{a}$ *y* $\mathfrak{b}$, $\mathfrak{a} + \mathfrak{b} = \mathfrak{b} + \mathfrak{a}$. En efecto, si $\mathfrak{a} = \mathfrak{b}$, es claro que $\mathfrak{a} + \mathfrak{b} = \mathfrak{a} + \mathfrak{a} = \mathfrak{b} + \mathfrak{a}$. De otro modo, podemos suponer que $\mathfrak{b} > \mathfrak{a}$. En tal caso, hay un número $\mathfrak{c}$ tal que $\mathfrak{b} = \mathfrak{a} + \mathfrak{c}$. Hay que demostrar, entonces, que $\mathfrak{a} + \mathfrak{a} + \mathfrak{c} = \mathfrak{a} + \mathfrak{c} + \mathfrak{a}$. Pero esto es obvio si $\mathfrak{a} + \mathfrak{c} = \mathfrak{c} + \mathfrak{a}$. Por lo tanto, como $\mathfrak{a} + \mathfrak{b} > \mathfrak{b} = \mathfrak{a} + \mathfrak{c}$, el teorema queda demostrado para *cualquier* par de números $\langle\mathfrak{a}, b\rangle$ con sólo que sea válido para cada par $\langle\mathfrak{a}, \mathfrak{c}\rangle$ tal que $\mathfrak{a} + \mathfrak{b} > \mathfrak{a} + \mathfrak{c}$.

Mediante el uso de variables que representan números cualesquiera y de razonamientos inductivos como el que acabo de parafrasear, Hilbert salta, más allá de toda "vivencia inmediata", a la consideración de lo que podemos llamar *vivencias posibles*—tomando, claro está, el epíteto 'posible' en un sentido bien amplio, no limitado por la cantidad de papel y tinta disponibles para anotar números ni por nuestra capacidad para reconocerlos y distinguirlos cuando constan de muchas copias de los objetos primarios 1 y +.[14] Consciente de ello, Hilbert declara que el "supuesto mínimo" arriba descrito no proviene de la experiencia ni del pensamiento (*Erfahrung und Denken*), sino de "una tercera fuente de conocimiento": una "disposición intuitiva a priori (*anschauliche Einstellung a priori*)" que constituye "la condición de la posibilidad de todo conocimiento" (1931, p. 486).

Esta intuición a priori hilbertiana —en contraste con la de Kant— es intrínsecamente finita. Su ámbito constituye el dominio propio del "punto de vista finito" o la "postura finita", el suelo firme en que Hilbert cimentará la matemática.[15] Por otra parte, la matemática como la conocemos no tiene

14 Si Hilbert se atuviera a estas limitaciones y $\mathfrak{b} = \mathfrak{a} + \mathfrak{c}$ fuese el número más grande que el cerebro humano y nuestros demás recursos nos permiten escribir y leer, no podría pretender que $\mathfrak{a} + \mathfrak{b} = \mathfrak{b} + \mathfrak{a}$, aunque hubiera demostrado paso a paso que $\mathfrak{a} + \mathfrak{c} = \mathfrak{c} + \mathfrak{a}$.

15 Cf. Hilbert 1931, p. 486: "Das Apriori ist dabei nichts mehr und nichts weniger als eine Grundeinstellung, die ich auch als die finite Einstellung bezeichnen möchte. (El a priori aquí no es ni más ni menos que una postura fundamental que quisiera también llamar *la postura finita*)" Al incluir en este ámbito de saber incuestionable el tipo de inferencia inductiva que se usa en la demostración citada, Hilbert parece darle la razón a Poincaré en su polémica contra el logicismo del Frege y Russell. Poincaré rechazó el intento de basar la matemática en la lógica porque la edificación rigurosa de la lógica

escrúpulos en hablar de lo infinito. Según Hilbert, al emplear los conceptos ‘todo *x*’ y ‘hay un *x*’ ya hemos rebasado “lo concretamente intuitivo y finito” (1923, en GA III, 181). Pero estos conceptos pueden eliminarse cuando se refieren a un agregado finito. “La aseveración de que *todos* los objetos de una colección finita abarcable dada tienen una cierta propiedad equivale lógicamente a la reunión de muchas aseveraciones particulares mediante la conjunción ‘y’ [ . . . ]. Del mismo modo, la aseveración de que en una colección finita *hay* un objeto con una propiedad equivale al enlace de aseveraciones particulares mediante la conjunción ‘o’” (GA III, 181). De esto se desprende inmediatamente que en cualquier colección finita, o bien todos los objetos tienen una cierta propiedad, o bien hay uno que no la tiene. Los matemáticos habitualmente dan por descontado sin más que este principio del “Tercero excluido” (*Tertium non datur*) vale también para las colecciones

---

requiere la inducción matemática (Poincaré 1894, en SH, pp. 38–42; 1905/1906, en CM, pp. 116, 128, 131). Hilbert ya lo había comprendido en 1904 cuando propugna una fundamentación *simultánea* de la lógica y de la aritmética (GG[7], p. 250; cf. 1922, en GA III, 174). Pero Hilbert advierte en más de una ocasión (1922, en GA III, 164; 1928, en GG[7], p. 301; 1928a, en GG[7], pp. 320s.) que el género de razonamiento inductivo que él acepta y practica sin discusión es muy distinto de la inducción matemática basada en el axioma **P5** de Peano (v*ide* p. 146). Cuando demostramos inductivamente que $\mathfrak{a} + \mathfrak{b} = \mathfrak{b} + \mathfrak{a}$ cualesquiera que sean los números (filas de signos) $\mathfrak{a}$ y $\mathfrak{b}$, discurrimos sobre la construcción de dos filas de signos de la forma general $1 + 1 + \ldots + 1$. El “paso inductivo” se reduce a ver que, dada una fila $\phi$ de este tipo, la fila $\phi + 1$ que se obtiene añadiendo la fila +1 a la derecha de $\phi$, es idéntica a la fila $1 + \phi$ que se obtiene al poner la fila 1+ a la izquierda de $\phi$. Del mismo género son los razonamientos por inducción sobre el número de paréntesis, o de conectivos, o de cuantificadores, etc. que se emplean —por ejemplo, en nuestro Apéndice IX— para establecer propiedades generales de las fórmulas de un cálculo lógico. Pero el axioma **P5** de Peano autoriza inferencias mucho más ambiciosas. Sea *n* un cierto número y sea $P(n)$ una aseveración de cualquier grado de complejidad referente a *n*. Sea $P(\mathfrak{x})$ la expresión que se obtiene al reemplazar en $P(n)$ todas las referencias a *n* por referencias al número indeterminado $\mathfrak{x}$. Entonces, **P5** nos permite inferir la validez de $P(\mathfrak{x})$, cualquiera que sea el número $\mathfrak{x}$, de las dos premisas siguientes: (i) $P(1)$ y (ii) cualquiera que sea $\mathfrak{x}$, $\neg P(\mathfrak{x}) \vee P(\mathfrak{x} + 1)$. La diferencia entre este caso y el anterior estriba en la complejidad de la aseveración demostrada: $P(n)$ puede referirse mediante cuantificadores existenciales y universales a toda suerte de números indeterminados. Por eso, Hilbert considera que la legitimidad de la inferencia inductiva basada en **P5** debe justificarse demostrando la consistencia de los axiomas de la aritmética, y que en esta demostración cabe emplear, sin circularidad, el otro género de razonamiento inductivo .

infinitas. Pero, “así como en el análisis está permitido transferir a sumas y productos infinitos los teoremas válidos para sumas y productos finitos sólo cuando esta inferencia está garantizada por una investigación especial de las condiciones de convergencia, así también, no podemos tratar las sumas y productos lógicos $A_1 \wedge A_2 \wedge A_3 \wedge \ldots$ y $A_1 \vee A_2 \vee A_3 \vee \ldots$ como si fueran finitos, a menos que la teoría de la prueba [ . . . ] autorice este tratamiento” (GA III, 182; uso nuestro signo ‘$\wedge$’ en vez del original ‘&’).

Hilbert quiere preservar intacto el discurso matemático sobre el infinito, el cual está ya vindicado a sus ojos por la ausencia de contradicciones en el análisis clásico y por el éxito de sus aplicaciones en la ciencia natural.[16] Pero quiere también establecer sin lugar a dudas y de una vez por todas que esa forma de discurso es legítima. Para eso, según él, hay que anclarlo firmemente en el dominio intuitivo de la “postura finita”: “El operar con lo infinito sólo puede asegurarse a través de lo finito” (1926, p. 190; cf. 1931, p. 488); “la libre manipulación y pleno control de lo transfinito tiene que alcanzarse en el terreno de lo finito” (1923, en GA III, 182). Para explicar lo que este proyecto significa, Hilbert evoca la sustitución de los infinitésimos de Newton y Leibniz por la “epsilóntica” de Cauchy y Weierstraß:[17]

> Así como en los procesos de paso al límite en el cálculo infinitesimal se ha podido mostrar que lo infinito en el sentido de lo infinitamente pe-

[16] “Todos los héroes del espíritu antes de Gauß, y también los que siguen a Gauß, Hermite y Jacobi hasta Poincaré, han empleado la inferencia transfinita en las formas más variadas y audaces, sin que nunca se haya manifestado ni la más leve discordancia. Por último, cuando pensamos en las aplicaciones y nos hacemos cargo de la riqueza en inferencias transfinitas del tipo más difícil y laborioso que albergan, por ejemplo, la teoría de la relatividad y la teoría cuántica y cómo, sin embargo, la naturaleza se ajusta precisamente a sus resultados —el rayo de luz de la estrella [curvado cuando pasa cerca del sol], Mercurio y los espectros más complicados aquí en la tierra y a una distancia de cientos de miles de años luz— ¿será posible en tales circunstancias que, a causa de los lindos ojos de Kronecker y unos cuantos filósofos disfrazados de matemáticos, por razones que además son completamente arbitrarias y ni siquiera se dejan formular con precisión, dudemos siquiera un instante de que la aplicación del Tertium non datur se justifica?” (Hilbert 1931, pp. 487s.).

[17] El término ‘epsilóntica’ alude a la práctica habitual en los cursos de análisis de utilizar la minúscula griega epsilon ($\varepsilon$) para referirse a una cantidad arbitrariamente pequeña. Por ejemplo, decimos que la función $f\colon \mathbb{R} \to \mathbb{R}$ es continua en el argumento $a$ si, para cualquier $\varepsilon > 0$, hay un $\delta > 0$ tal que $|f(x) - f(a)| < \varepsilon$ si $|x - a| < \delta$.

> queño e infinitamente grande no era más que *una manera de hablar*, así también lo infinito en el sentido de la colección infinita, como aún ahora se nos presenta en los modos de inferencia, tiene que reconocerse como algo *meramente aparente*. Y así como el operar con lo infinitamente pequeño fue reemplazado con procesos en el dominio finito que efectúan lo mismo y llevan a las mismas elegantes relaciones formales, así también en general hay que reemplazar los modos de inferencia que envuelven lo infinito con procesos finitos que efectúan lo mismo, es decir, que hacen posibles las mismas demostraciones y los mismos métodos para obtener fórmulas y teoremas.
>
> (Hilbert 1926, p. 162; cursiva mía)[18]

Otra analogía favorita de Hilbert son los "elementos ideales" introducidos en diversas teorías matemáticas para redondearlas o simplificarlas.

> Así como $i = \sqrt{-1}$ se introdujo para mantener en pie en su forma más simple las leyes del álgebra —por ejemplo, las que se refieren a la existencia y número de las raíces de una ecuación—; así como los factores ideales se introdujeron para preservar también entre los números enteros algebraicos las leyes simples de divisibilidad —y así, por ejemplo, introducimos un divisor común de los números 2 y $(1 + \sqrt{-5})$, aunque no tienen uno real—; así también tenemos que *adjuntar* (*adjungieren*) *los enunciados ideales a los enunciados finitos*, para preservar las reglas formales simples de la logica aristotélica usual.
>
> (Hilbert 1926, p. 174; cf. 1928 en GG[7], pp. 298s.)

El origen y la índole de tales enunciados ideales puede ilustrarse con el ejemplo del álgebra y su método de cálculo literal. Consideremos una vez más el teorema arriba demostrado: $\mathfrak{a} + \mathfrak{b} = \mathfrak{b} + \mathfrak{a}$, cualesquiera que sean los números $\mathfrak{a}$ y $\mathfrak{b}$. Aquí las letras góticas $\mathfrak{a}$ y $\mathfrak{b}$ son, como dijimos, signos con significado: cada una representa un número indeterminado cualquiera. Pero en la práctica ordinaria del álgebra este significado se olvida: el cálculo literal maneja las letras como objetos primarios que no remiten a nada más allá, "signos sin significado" como los signos 1 y + de la aritmética finita de

18 A este mismo contexto pertenece la frase de Hilbert en que se inspira el título de este libro: "Del paraíso que Cantor creó para nosotros, nadie podrá expulsarnos" (1926, p. 170). Las expresiones que destaqué en el pasaje arriba citado indican que Hilbert concibe este paraíso como un *set* cinematográfico: el Jardín de las Delicias pintado sobre papel.

Hilbert. Por eso Hilbert puede decir (1926, p. 175) que la fórmula algebraica

$$a + b = b + a$$

no constituye "una comunicación directa de algo sustantivo"[19] sino que es sólo una "figura formal", cuya relación con los enunciados finitos

$$3 + 2 = 2 + 3$$

$$5 + 7 = 7 + 5$$

consiste en que éstos pueden *inferirse* de aquélla en virtud de los axiomas y reglas de inferencia comúnmente aceptados. Generalizando este enfoque, la matemática puede verse como una provisión de fórmulas, que en parte expresan enunciados finitos de contenido sustantivo y en parte *no significan nada*. Estas últimas son los enunciados ideales.

> Como los enunciados ideales —esto es, las fórmulas que no expresan aseveraciones finitas— no significan nada, las operaciones lógicas no pueden aplicárseles de un modo sustantivo como a los enunciados finitos. Es necesario entonces formalizar también las operaciones lógicas y las demostraciones matemáticas mismas.
>
> (Hilbert 1926, p. 176)[20]

[19] Como expliqué en la p. 123, nota 13, uso 'sustantivo' para traducir el término hilbertiano 'inhaltlich'.

[20] Hilbert prosigue diciendo que, gracias a esa "armonía preestablecida que tan a menudo se observa en la historia evolutiva de la ciencia", dicha formalización está ya muy adelantada por obra de los creadores del *cálculo lógico*.

> Los símbolos del cálculo lógico se introdujeron originalmente sólo para la comunicación; pero es consecuente que ahora neguemos todo significado a los signos lógicos, igual que a los matemáticos, y declaremos que también las fórmulas del cálculo lógico no significan nada de por sí, sino que son enunciados ideales. En el cálculo lógico poseemos un lenguaje de signos que es capaz de abarcar en fórmulas los teoremas matemáticos y de expresar la inferencia lógica mediante procesos formales. Igual que en el tránsito de la aritmética sustantiva (*inhaltliche Zahlenlehre*) al álgebra formal, contemplamos los signos y símbolos de operaciones del cálculo lógico prescindiendo de su significado sustantivo. Así obtenemos finalmente, en lugar de la ciencia matemática sustantiva que se trasmite mediante el lenguaje ordinario, una provisión de fórmulas con signos matemáticos y lógi-

Mediante la formalización de las demostraciones, los enunciados ideales y los enunciados finitos se entrelazan en un sistema formal de matemáticas. ¿En qué relación quedan éstos con aquéllos dentro de dicho sistema? Para arrojar luz sobre esta cuestión me referiré a un famoso ejemplo de "elementos ideales" que Hilbert menciona a veces: los "puntos infinitamente lejanos" de la geometría proyectiva (cf. Hilbert 1923, p. 187; 1926, p. 166). Aunque la invención de este recurso puede atribuirse a Kepler y fue utilizado por Desargues y Pascal en el siglo XVII, el auge de la geometría proyectiva comienza en 1822 con la publicación del *Traité des propriétés projectives des figures* de Poncelet. Sesenta años más tarde, Pasch (1882) la llama simplemente 'geometría moderna'. El giro característico que conduce de la geometría tradicional a la proyectiva puede describirse sumariamente así: Consíderese un plano euclidiano; a cada recta le agregaremos un objeto que llamamos 'el *punto en el infinito*' de esa recta. Con este agregado, habremos formado el *plano proyectivo* si sobreentendemos que (i) si $p$ es el punto en el infinito de una recta cualquiera $\lambda$ y $a$, $b$ y $c$ son tres puntos de $\lambda$ tales que $b$ está entre $a$ y $c$, entonces $a$ está entre $p$ y $b$, y $c$ está entre $b$ y $p$ (en virtud de esto es claro que el conjunto de *todos* los puntos de $\lambda$ tiene el mismo tipo de orden que los puntos de un círculo); (ii) el punto en el infinito de $\lambda$ es idéntico al punto en el infinito de cualquier recta paralela a $\lambda$, de modo que todas las rectas de una familia de paralelas se intersectan en su respectivo punto en el infinito; (iii) los puntos en el infinito pertenecientes a cada familia de paralelas forman la 'recta en el infinito' del plano proyectivo; el orden (cíclico) de los puntos de la recta en el infinito se deriva fácilmente del orden en que las rectas ordinarias que pasan por un mis-

---

cos, enfiladas unas tras otras según ciertas reglas. A los axiomas matemáticos corresponden algunas de las fórmulas, y a la inferencia sustantiva corresponden las reglas de acuerdo con las cuales se suceden las fórmulas. De este modo, la inferencia sustantiva es reemplazada por un operar extrínseco (*ein äußeres Handeln*) conforme a reglas y se consuma rigurosamente el tránsito del tratamiento ingenuo al tratamiento formal con respecto a los axiomas mismos —que en un principio se propusieron ingenuamente como verdades fundamentales, pero que la axiomática moderna considera hace tiempo como meras combinaciones de conceptos— y también con respecto al cálculo lógico, que originalmente no pretendía ser sino un lenguaje más.

(1926, pp. 176s.; cf. GA III, 165, 179; GG[7], pp. 298s.).

mo punto ordinario *q* intersectan un círculo cualquiera con centro en *q*, ya que cada punto de la recta en el infinito pertenece a una y sólo una de esas rectas ordinarias. Los puntos en el infinito han solido considerarse como meras ficciones, pero también se los ha identificado, por ejemplo, con la dirección de la recta respectiva: no cabe duda que cada recta del plano euclidiano tiene una dirección y que tal dirección es común a todas las rectas de una familia de paralelas. La innovación de la geometría proyectiva consiste entonces en definir convencionalmente una relación de orden —descrita en (i) y (ii)— entre los puntos de una recta y su dirección (como se define en derecho una relación legal de parentesco entre los consanguíneos de una persona y los consanguíneos de su cónyuge). Más satisfactorio aún es concebir a los elementos del plano proyectivo como objetos cualesquiera, cuyas propiedades y relaciones se especifican mediante postulados, siguiendo el ejemplo de Hilbert (1899, 1900). Visto de este modo, el plano proyectivo resulta ser no una ficción matemática arbitraria, ni tampoco un monstruo quimérico generado asimilando las direcciones de rectas a los puntos del plano, sino una estructura abstracta realizable en objetos familiares. Concretamente, el haz de todas las rectas que pasan por un punto *p* del espacio ordinario constituye una realización del plano proyectivo si convenimos en llamarlas 'puntos', llamamos 'recta' a cada plano determinado por dos de ellas, y designamos un plano cualquiera como 'la recta en el infinito' (lo arbitrario de esta selección refleja la circunstancia de que en la teoría del plano proyectivo no hay, en efecto, ninguna diferencia entre la recta en el infinito y las otras rectas). Si eliminamos del haz el plano elegido para desempeñar este papel, resta un sistema de 'puntos' y 'rectas' que llamaré los 'elementos finitos' del plano proyectivo. Dicho sistema reproduce las relaciones de incidencia y orden entre sus homónimos en el plano euclidiano. Por lo tanto, cualquier teorema de la geometría proyectiva plana concerniente a relaciones de incidencia u orden entre elementos finitos expresa una verdad de la geometría euclidiana plana.

Aplicando la analogía de la geometría proyectiva al sistema de la matemática formal concluimos que *dentro de éste* los enunciados finitos no se distinguen de los enunciados ideales: unos y otros están a la par. Por lo tanto, desde un punto de vista inmanente al sistema, aquéllos no tienen ni una brizna más de significado que éstos. La diferencia, extrínseca al sistema formal, viene de que los enunciados finitos tienen —y los ideales no— una traducción dotada de sentido al lenguaje informal e intuitivo de la práctica

matemática ordinaria y de la experiencia. La utilidad de los enunciados ideales reside esencialmente en esto: en el sistema enriquecido con ellos es más fácil validar muchos enunciados finitos cuya traducción dotada de sentido es verdadera. Como es obvio, tal utilidad se perdería si el sistema formal permitiera validar algún enunciado finito cuya traducción fuese falsa. En otras palabras: para ser útil, el sistema de la matemática formal debe constituir una extensión conservadora del sistema de los enunciados finitos. Evidentemente, esto es así sólo si la matemática formal es consistente, esto es, si no todo enunciado formulable en el sistema es a la vez validable en el sistema. Introduciendo más precisión en el planteamiento del problema se puede probar que esta condición necesaria también es suficiente, esto es, que si el sistema formal es consistente todo enunciado finito validable en él es verdadero (Smorynski 1977, pp. 846-47; cf. pp. 823-24). La matemática clásica quedará, pues, vindicada si se establece la consistencia de su versión formalizada. Con este fin, Hilbert hace del sistema mismo de la matemática formal un objeto de investigación matemática en el sentido tradicional. La investigación puede conducirse sin infringir las limitaciones impuestas a la matemática sustantiva, ya que los enunciados del sistema formal son secuencias finitas de signos tomados de un alfabeto finito y las derivaciones que validan un enunciado dado a partir de otros ya validados son secuencias finitas de enunciados, y ambos géneros de secuencias obedecen a reglas perspicuas. Hilbert llama 'metamatemática' (*Metamathematik*) o también 'teoría de la prueba' (*Beweistheorie*) a la disciplina matemática —una especie de análisis combinatorio— encargada de tal investigación. Ella constituye un discurso *con sentido*, un razonar *sustantivo* acerca de los objetos finitos —signos, fórmulas, pruebas— que componen la matemática formalizada. Se trata, en suma, de mostrar que un enunciado cualquiera, por ejemplo '$1 \neq 1$', no se puede validar derivándolo en el sistema formal. ¿Quién podría dudar de que esta tarea es abordable de un modo puramente intuitivo? Hilbert la compara con la demostración, conocida desde la antigüedad, de que $\sqrt{2}$ es irracional. Para establecerlo demostramos intuitivamente que no puede haber dos números $\mathfrak{a}$ y $\mathfrak{b}$ tales que $\mathfrak{a}^2 = 2\mathfrak{b}^2$, o sea, que no pueden existir dos numerales constituidos del modo descrito.[21] "Análogamente, nos corresponde mostrar

[21] La demostración, trasmitida en uno de los apéndices apócrifos de los *Elementos* de Euclides, puede formularse así: Si $\sqrt{2}$ no es irracional, hay dos números $\mathfrak{a}$ y $\mathfrak{b}$, primos

que no puede haber una prueba constituida de cierto modo. Pero una prueba formalizada, lo mismo que un numeral, es un objeto concreto y abarcable con la mirada. Es comunicable de principio a fin. También la requerida constitución de la fórmula final, a saber, que lea '$1 \neq 1$', es una propiedad concretamente constatable de la prueba. Por lo tanto, esta demostración de hecho puede suministrarse, y con ella ganamos el derecho a introducir nuestros enunciados ideales" (1926, p. 179; cf. 1928, en GG$^7$, p. 300).[22]

### *2.9.4 La investigación de la consistencia de la matemática formalizada, a la luz del descubrimiento de Gödel*

El programa de Hilbert propone, pues, probar la consistencia del sistema formal de la matemática clásica con medios finitos razonando sustantivamente sobre sus enunciados y derivaciones, considerados como combinaciones de signos sin sentido. Esta empresa, al parecer tan factible, ha sido puesta en cuestión en vista de los resultados que Gödel anunció a fines de 1930 y demostró rigurosamente en su artículo de 1931, que examinaremos en detalle en el Capítulo 2.10. Me parece oportuno explicar aquí sumariamente en qué sentido preciso el programa de Hilbert queda afectado por dichos resultados. Así podré afinar un poco más mi descripción de este programa y tam-

---

entre sí, tales que $\mathfrak{a}/\mathfrak{b} = \sqrt{2}$. En tal caso, $\mathfrak{a}^2 = 2\mathfrak{b}^2$. Por lo tanto, $\mathfrak{a}^2$ es un número par. Pero un número par no puede ser el cuadrado de un número impar (si $\mathfrak{x} = 2\mathfrak{n} + 1$, entonces $\mathfrak{x}^2 = 4\mathfrak{n}^2 + 2\mathfrak{n} + 1$, que es impar). Por lo tanto, $\mathfrak{a}$ es par: $\mathfrak{a} = 2\mathfrak{c}$. Por lo tanto, $2\mathfrak{b}^2 = \mathfrak{a}^2 = 4\mathfrak{c}^2$ y $\mathfrak{b}^2 = 2\mathfrak{c}^2$. Por lo tanto, también $\mathfrak{b}$ es par. Por lo tanto, $\mathfrak{a}$ y $\mathfrak{b}$ no son primos entre sí. ¡Contradicción! Por lo tanto, $\sqrt{2}$ es irracional.

22 El optimismo de Hilbert sería más justificado si la matemática formalizada fuese inconsistente y se tratase de mostrarlo mediante razonamientos sustantivos sobre objetos finitos. Para eso sí que nos ayudaría la índole "concretamente constatable" de la fórmula "$1 \neq 1$". Bastaría construir una prueba formal que desemboque en esta fórmula. Pero la tarea de la metamatemática hilbertiana consiste en mostrar mediante tales razonamientos que la fórmula "$1 \neq 1$" *no puede aparecer nunca* en la última línea de una prueba de la matemática formalizada; y no veo bien por qué Hilbert daba tan fácilmente por descontado que esta aseveración referente a *todas* las *infinitas* pruebas en principio enunciables podría demostrarse con medios finitos.

bién, espero, motivar al lector para el esfuerzo de atención que requiere el estudio de la obra de Gödel.

Gödel considera una formalización de la aritmética elemental ajustada al prototipo de *Principia Mathematica* "y sistemas afines". Las investigaciones de Church, Turing y otros (Capítulo 2.11) han permitido determinar con precisión los requisitos que tiene que cumplir un sistema formal para que se le apliquen los resultados de Gödel y sea, en este sentido, afín a *Principia Mathematica*. Se trata sencillamente de esto: que existan procedimientos de cómputo para decidir, en un número finito de operaciones, si un objeto dado es o no un signo del sistema, si una secuencia de tales signos es o no un enunciado del sistema y si una secuencia de tales enunciados es o no una derivación —una "prueba"— del sistema. No me cabe ninguna duda de que los sistemas formales estudiados en detalle en la década de los 20 por los seguidores de Hilbert —Ackermann, Herbrand, von Neumann— llenaban los requisitos indicados. Por otra parte, en un escrito aparecido el mismo año que el artículo de Gödel, Hilbert propone adoptar en la matemática formalizada una regla de inferencia que infringe esos requisitos, por cuanto autoriza derivaciones con una lista infinita de premisas. La referida regla dice así:

> Si se ha comprobado que la fórmula $\mathfrak{A}(\mathfrak{z})$ —donde $\mathfrak{z}$ es un numeral dado— es siempre [esto es, para cualquier numeral—R.T.] una fórmula numérica correcta, puede usarse como premisa la fórmula $\forall x \mathfrak{A}(x)$.[23]

Hilbert agrega que la nueva regla de inferencia es "finita", lo que indica —me parece— que entendía esta palabra clave en una acepción muy personal. Los sistemas formales que Hilbert y Bernays estudian en el *magnum opus* publicado en 1934 y 1939 no incluyen esta regla, y son todos afines a *Principia Mathematica* en el sentido explicado. Ello no obstante, no debemos perder de vista que Hilbert, aún antes de estudiar los resultados de Gödel, estaba dispuesto a favorecer un método de formalización que los evade.

[23] Hilbert 1931, p. 491. La regla citada suele llamarse "la regla ω" (cf. Isaacson 1992). Hilbert comenta que la "oración (*Aussage*)" $\forall x \mathfrak{A}(x)$ va mucho más lejos que la "fórmula (*Formel*)" $\mathfrak{A}(\mathfrak{z})$, con $\mathfrak{z}$ un numeral cualquiera, puesto que aquélla autoriza a inferir toda oración que se obtenga reemplazando en $\mathfrak{A}(x)$ la variable libre $x$ por *cualquier término* que denote un número (no sólo por *cualquier numeral*).

Sea $\mathcal{S}$ un sistema formal de aritmética del género descrito. Como luego veremos, Gödel ideó un método para representar mediante números los signos, enunciados y derivaciones de $\mathcal{S}$. Las propiedades y relaciones sintácticas de dichos objetos de $\mathcal{S}$ vienen a quedar representadas por propiedades y relaciones aritméticas de los números correspondientes. La investigación metamatemática de $\mathcal{S}$ con medios finitos se refleja entonces exactamente en una investigación combinatoria sobre dichos números, sus propiedades y relaciones, cuyos razonamientos y resultados se dejan por cierto formalizar en $\mathcal{S}$. En particular, puede definirse un predicado aritmético $P$ atribuible a un número si y sólo si ese número representa una fórmula derivable en $\mathcal{S}$. Entonces, si $a$ es el número representativo de la fórmula de $\mathcal{S}$ que expresa el enunciado '$1 \neq 1$', el enunciado que atribuye el predicado $P$ al número $a$ es verdadero si y sólo si $\mathcal{S}$ es inconsistente y la negación de este enunciado es verdadera si y sólo si $\mathcal{S}$ es consistente. Y habrá muchos otros enunciados que, como éste, representan la consistencia de $\mathcal{S}$. Como son enunciados aritméticos, cada uno de ellos se expresa formalmente mediante una fórmula de $\mathcal{S}$. Sea 'Consis $\mathcal{S}$' la fórmula correspondiente a cualquiera de ellos. Gödel demostró que Consis $\mathcal{S}$ *no puede derivarse en* $\mathcal{S}$ *a menos que* $\mathcal{S}$ *sea inconsistente*. Este es el resultado que, en opinión de la mayoría de los matemáticos, dio un golpe de muerte al programa de Hilbert.[24] A la luz de las explicaciones precedentes, no cabe duda de que este diagnóstico es acertado si la investigación metamatemática se deja representar, como se ha dicho, por una

[24] El pasaje siguiente expresa la opinión del grupo Bourbaki:

> Es en efecto en lo relativo a la cuestión de la no-contradicción de las teorías matemáticas —el origen y el corazón mismo de la metamatemática— que los resultados se han revelado más defraudantes. Durante los años 1920-1930 Hilbert y su escuela desarrollaron métodos nuevos para abordar estos problemas; tras haber demostrado la no-contradicción de formalismos parciales, que cubrían una parte de la aritmética, creían estar a un paso de la meta y demostrar no sólo la no-contradicción de la aritmética sino también la de la teoría de conjuntos, cuando Gödel, apoyándose en la no-categoricidad de la aritmética, dedujo la imposibilidad de demostrar mediante los "procedimientos finitos" de Hilbert, la no-contradicción de cualquier teoria $\mathcal{T}$ que contenga a la aritmética.
>
> Sin embargo, el teorema de Gödel no cierra totalmente la puerta a las tentativas de demostrar la no-contradicción, con tal que se abandonen (al menos parcialmente) las restricciones de Hilbert concernientes a los "procedimientos finitos". Es así como Gentzen en 1936 logró demostrar la no-contradicción de la arit-

investigación aritmética formalizable en $\mathcal{S}$; pero el diagnóstico no tiene base si la investigación metamatemática dispone de recursos que no pueden representarse así. Ahora bien, parecería que los recursos y procedimientos representables en la aritmética elemental y, por ende, en $\mathcal{S}$ abarcan todo lo que una persona sensata pudiera describir como 'métodos finitos'. Por otra parte, como veremos en el Capítulo 2.12, Gerhard Gentzen (1935, 1938) demostró la consistencia de un sistema formal de aritmética del tipo estudiado por Gödel empleando inducción transfinita hasta el ordinal numerable que Cantor llamó $\varepsilon_0$ (el primer ordinal $\xi$, tal que $\omega^\xi = \xi$.). Hilbert y Bernays reconocieron que los métodos de Gentzen tenían cabida dentro del programa de Hilbert y eran compatibles con su punto de vista finitista. Para ello tuvieron que "ampliar el marco de los modos de inferencia sustantivos admitidos en la teoría de la prueba" (1939, p. vii; cito el original en la p. 421, nota 1), pero esto no les causó, al parecer, ningún escrúpulo. Resta, sin embargo, una duda: si el programa de Hilbert acaba recurriendo al transfinito, ¿por qué tantos melindres y reservas ante el paraíso heredado de Cantor? ¿por qué no instalarse en él, alegremente, de una vez por todas?

---

> mética formalizada utilizando "intuitivamente" la inducción transfinita hasta el ordinal numerable $\varepsilon_0$. El valor de "certeza" atribuible a tal razonamiento es sin duda menos probatorio que en el caso de los que satisfacen a las exigencias iniciales de Hilbert y es esencialmente cosa de la psicología personal de cada matemático.
>
> (Bourbaki 1970, E.IV.75-76)

Últimamente algunos filósofos han alegado que el citado descubrimiento de Gödel no constituye un obstáculo para el programa de Hilbert (Webb 1980, Detlefsen 1986, 1990). Pero hasta la fecha no han establecido que la consistencia de la aritmética se pueda demostrar con recursos más débiles que los utilizados por Gentzen. Me referiré nuevamente a Detlefsen una vez que hayamos estudiado los resultados de Gödel (p. 358, nota 48).

## 2.10 LOS TEOREMAS DE INCOMPLETUD DE GÖDEL

### *2.10.1 Preliminares*

En este capítulo estudiaremos el escrito de Gödel "Sobre proposiciones formalmente indecidibles de *Principia Mathematica* y sistemas afines" (1931). Su enorme resonancia histórica responde, en parte, a sus sorprendentes resultados, pero también a los métodos de que el autor se vale para establecerlos. Los métodos de Gödel sientan las bases para una nueva disciplina matemática, la teoría de las funciones recursivas o "computables". El gran impacto de esta disciplina sobre la vida diaria viene de que ella es el núcleo teórico en torno al cual se desarrolla, en la segunda mitad del siglo XX, la llamada "informática" y la industria de las computadoras. Mas desde un punto de vista estrictamente científico y filosófico su significación radica en que hay buenas razones para pensar que el concepto preciso de función computable, definido de varias maneras equivalentes por sus cultivadores, abarca todas las funciones numéricas (con argumentos y valores en $\mathbb{N}$) cuyo valor, para cada argumento dado, puede calcularse mediante un algoritmo, en el sentido intuitivo corriente.[1] En el Capítulo 2.7 (p. 247), apelé a la idea intuitiva de algoritmo para caracterizar los sistemas formales o "cálculos" a que Hilbert y su escuela querían reducir la matemática clásica. Si esa idea intuitiva —y vaga— se reemplaza con el concepto preciso que es el tema de la referida disciplina, los resultados obtenidos por Gödel limitan de un modo claro y terminante el alcance de cualquier representación de teorías matemáticas mediante tales cálculos y la utilidad de los mismos para la fundamentación de las matemáticas.

[1] Como bien señala Post, "si la función recursiva general [una de esas nociones precisas de función computable—R.T.] es el equivalente formal de la calculabilidad efectiva, su formulación podría tener en la historia de la matemática combinatoria una importancia superada sólo por la formulación del concepto de número natural" (1944, p. 315).

Gödel concentra su atención en cierto cálculo *P*, que se parece mucho al utilizado en *Principia Mathematica*.[2] Pero sus razonamientos se basan sólo en ciertas características muy generales de *P*, y sus conclusiones valen, por ende, para todo cálculo que las comparta. Desde luego, es claro que tanto el cálculo de *Principia Mathematica* como los utilizados por Ackermann, von Neumann y Herbrand en sus investigaciones metamatemáticas se hallan en ese caso, y que costaría muy poco adaptar la argumentación de Gödel a cualquiera de ellos.[3] Mas para enunciar los teoremas de Gödel con toda generalidad, habría que decir exactamente en qué consisten las características en cuestión. Así se deslindaría la clase de los "sistemas afines" a *Principia Mathematica* mencionados el título de Gödel 1931. Ahora bien, dichas características se reducen esencialmente a dos: (i) cada cálculo o sistema de esa clase provee una representación formal de la aritmética elemental; y (ii) la sintaxis del cálculo puede a su vez representarse mediante números naturales y sus propiedades y relaciones. Luego explicaré en detalle el significado de (i) y (ii) a propósito del cálculo *P* de Gödel.[4] En todo caso, la idea de una representación formal de la aritmética no tomará por sorpresa al lector de los capítulos anteriores (especialmente, el 2.6). En cambio, la idea de una representación *numérica* de la sintaxis del cálculo es una de las contribuciones originales de Gödel, de una sencillez genial. Como veremos luego, la representación numérica requerida debe cumplir con los siguientes requisitos: (*a*) que cada objeto del cálculo, esto es, cada signo, cada fila de signos y cada lista de tales filas de signos, tenga un número de identidad exclusivo; (*b*) que, dado un número natural *n*, sea posible establecer, mediante un número finito de operaciones aritméticas elementales, si *n* es o no el número

2 En la comunicación preliminar de sus resultados a la Academia de Viena, Gödel (1930b) dice haberlos establecido para el caso del cálculo de *Principia Mathematica* "con Axioma de Reducibilidad o sin Teoría Ramificada de los Tipos" (CW, I, 140 n.1). De hecho, su cálculo *P* se ajusta a la segunda alternativa: es un cálculo predicativo de orden ω, conforme a la teoría simple de los tipos.

3 En cambio, como indiqué en la p. 317, la argumentación de Gödel no es aplicable a un cálculo cuyas reglas de inferencia incluyan la regla ω, allí citada, que Hilbert (1931, p. 491) propuso después que le comunicaron los resultados de Gödel.

4 El Apéndice XVI contiene una formulación precisa de las condiciones (i) y (ii), aplicable a cualquier cálculo. Ella proviene de una obra reciente (Smullyan 1992), donde se la utiliza en la presentación de la "forma abstracta" del primer teorema de incompletud de Gödel resumida en dicho Apéndice.

de identidad de una *prueba* (esto es, de una lista de fórmulas del cálculo, cada una de las cuales es un axioma del cálculo o se deduce, conforme a las reglas de inferencia del cálculo, de una o más fórmulas que la preceden en la lista) y, en caso afirmativo, (*c*) que sea posible reconstruir las fórmulas de que consta la prueba correspondiente. Gracias a esta doble representación —de la aritmética en el cálculo y del cálculo en la aritmética— Gödel puede cerrar el círculo de espejos en que descansa su Teorema VI, el "primer teorema de incompletud". En la Sección 2.10.2 describiré, paso a paso, la compleja argumentación que culmina en ese teorema. Pero antes conviene delinearla a grandes rasgos, para saber adónde lleva, y sobre todo para captar con un sólo golpe de vista la función que desempeñan en ella las características (i) y (ii). El bosquejo siguiente es una paráfrasis del ofrecido por el mismo Gödel (1931, § 1).

Gödel supone que toda oración deducible de los axiomas del cálculo considerado —esto es, toda oración que ocupa el último lugar de una lista que constituye una prueba— expresa una aseveración verdadera. Este supuesto es muy razonable, pues mal podría pretenderse formalizar la aritmética en un cálculo que genera teoremas falsos. Con todo, Gödel lo adopta sólo para facilitar la presentación del bosquejo preliminar: la demostración rigurosa del Teorema VI en el § 2 depende de supuestos puramente sintácticos, y no invoca el concepto de verdad. Para decir *en castellano* que $\phi$ es una oración deducible de los axiomas del cálculo, escribiremos $\vdash\phi$;[5] asimismo, la abreviatura $\nvdash\phi$ expresará que $\phi$ no es una oración deducible de los axiomas del cálculo. Supondremos que el cálculo contiene categorías sintácticas homólogas a las constantes y variables de nuestros cálculos predicativos (Apéndice IX).[6] Para que pueda representar formalmente la aritmética, el cálculo debe contener un conectivo —primitivo o introducido por definición— que prefijado a una oración verdadera genere una oración falsa y prefijado a una falsa genere una verdadera. Simbolizaré este conectivo *del cálculo* con el signo

5 Gödel usa la abreviatura *Bew*, del alemán 'beweisbar' ('demostrable'). Sus traductores al inglés y al castellano han preservado esta abreviatura, aunque en estos idiomas, a diferencia del original, ella obstaculiza la lectura: no sólo no evoca ninguna palabra pertinente, sino que ni siquiera se la puede pronunciar con facilidad.

6 Este requisito afecta, sin duda, la generalidad del argumento. Como puede verse en el Apéndice XVI, de hecho no se lo necesita para demostrar la "forma abstracta" del primer teorema de Gödel presentada allí.

russelliano de negación ~.[7] Diré que la oración φ es *demostrable* si es deducible de los axiomas, esto es, si $\vdash\phi$; y que es *refutable* si su negación es demostrable, esto es, si $\vdash\sim\phi$. Se trata de probar que hay en el cálculo por lo menos una oración que no es demostrable ni refutable.

Sea $\mathscr{F}_1$ la clase de las fórmulas que contienen una y sólo una variable libre. Si α es una fórmula de esa clase, designamos con $[\alpha;n]$ a la fórmula —en efecto, una oración— obtenida al sustituir la única variable libre de α (en todas sus posiciones) por el nombre del número *n*. Suponemos $\mathscr{F}_1$ ordenada según el orden de magnitud creciente de los números de identidad de sus fórmulas. Sea $R_n$ la *n*-ésima fórmula de $\mathscr{F}_1$ y *K* el conjunto de todos los números *n* tales que $\nvdash[R_n;n]$.[8] Pertenecer a *K* es una propiedad *de números* que se puede representar en el cálculo, si éste posee la característica (i). Tiene que haber, pues, una fórmula κ de la clase $\mathscr{F}_1$ que se transforma en una oración verdadera cuando su única variable libre es reemplazada por el nombre de un número que pertenece a *K* y en una oración falsa cuando esa variable es reemplazada por el nombre de un número que no pertenece a *K*. Esa fórmula tiene un número de identidad, digamos, el *q*-ésimo en el orden inducido en $\mathscr{F}_1$. La fórmula κ es, entonces, precisamente la fórmula que hemos convenido en llamar $R_q$. Consideremos la oración $[R_q;q]$. Esta oración no es demostrable ni refutable en el cálculo. En efecto, supongamos que $\vdash[R_q;q]$. Entonces $[R_q;q]$ es verdadera. Pero $[R_q;q]$ es la oración obtenida al reemplazar la variable libre en $R_q$ por el nombre del número *q*. Dicha oración es verdadera sí y sólo si $q \in K$, esto es, si y sólo si $\nvdash[R_q;q]$. Por lo tanto, $[R_q;q]$ no es demostrable. Supongamos, entonces, que su negación lo es, esto es, que $\vdash\sim[R_q;q]$. Entonces $[R_q;q]$ es falsa. Pero $[R_q;q]$ es la oración obtenida al reemplazar la variable libre en $R_q$ por el nombre del número *q*; dicha oración es falsa sí y sólo si $q \notin K$, esto es, si y sólo si $\vdash[R_q;q]$. Así pues, $\sim[R_q;q]$ sólo es demostrable si también lo es $[R_q;q]$, o sea, si el cálculo es inconsistente. Pero esto sería incompatible con el supuesto inicial de

7 Adopto esta convención porque en toda este capítulo utilizaré *nuestros* símbolos logicos como abreviaturas del discurso metalógico que conducimos en castellano. Imito así la práctica de Gödel, que destinó a tal propósito los símbolos de Hilbert y Ackermann 1928.

8 En otras palabras, el número *n* pertenece al conjunto *K* si y sólo si la oración obtenida al sustituir la variable libre en la fórmula $R_n$ por el nombre de *n* no es deducible de los axiomas de *P*.

que ninguna oración falsa es deducible de los axiomas.

El argumento precedente muestra que un cálculo que posee las características (i) y (ii) y en que toda expresión demostrable es verdadera contiene necesariamente una oración que no es demostrable ni refutable. En otras palabras, cualquier cálculo de la clase indicada es incompleto en el sentido de Post. La oración $[R_q;q]$, que lo atestigua, es verdadera si y sólo si el número $q$ pertenece al conjunto numérico

$$K = \{n\text{:}[R_n;n] \text{ no es deducible de los axiomas}\}.$$

Tenemos, entonces, que $[R_q;q]$ es verdadera, puesto que $[R_q;q]$, en efecto, no es demostrable. Por otra parte, como $[R_q;q]$ no es demostrable, si agregamos a la lista de axiomas la oración falsa $\sim[R_q;q]$, el sistema así formado es consistente. Por lo tanto, la verdad matemática no puede simplemente equipararse a la deducibilidad en un sistema consistente, como pretendía Hilbert.[9]

El razonamiento de Gödel ilustra el "método diagonal" descrito en la p. 43, nota 11. Como se recordará, dicho método se utiliza asimismo en la demostración del Teorema de Cantor y en la construcción de la paradoja de Russell. Gödel subraya la analogía entre su razonamiento y las paradojas de Richard y del mentiroso. Pero también la diferencia es obvia: "la fórmula φ es verdadera si y sólo si la negación de φ es verdadera" es una contradicción; pero "la fórmula ϕ es demostrable si y sólo si la negación de φ es demostrable" significa únicamente que la fórmula ϕ no puede deducirse de los axiomas convenidos aplicando las reglas de inferencia adoptadas (a menos que esos axiomas y reglas determinen un sistema deductivo inconsistente). Aunque defraudante para quien haya puesto su fe en los poderes del cálculo, esta conclusión no tiene nada de paradójica. De hecho, no debiera sorprendernos, puesto que, como se ha visto en los capítulos anteriores, el uso de cálculos lógicos implica un sacrificio feroz de la inteligencia en aras de la seguridad.

9 Recuérdese el pasaje de la carta de Hilbert a Frege del 29 de diciembre de 1899 (Frege, KS, p. 411), citado en p. 118.

### *2.10.2 La incompletud de la aritmética*

Concluido el bosquejo preliminar, Gödel emprende una demostración rigurosa de sus resultados, aplicable a cierto cálculo *P* bajo hipótesis sintácticas precisas relativas a su consistencia. El § 2 (1931, pp. 176-191), que concluye con el primer "teorema de incompletud" (Teorema VI), es una de las cumbres de la historia de la lógica, y creo que el lector debe darse el trabajo de leerlo. (Hay una buena traducción castellana en Gödel, OC). La explicación siguiente debiera facilitar, mas no sustituir esa lectura. Omitiré, pues, algunos detalles. Distinguiré cuidadosamente los pasos sucesivos del razonamiento, señalando la función de cada uno; pero descifraré, a modo de ejemplo, sólo una parte de las definiciones abreviadas de Gödel, dejando al lector la tarea, tediosa pero necesaria, de explicitar todas las otras.

1º El primer paso es la descripción de la sintaxis del cálculo *P*. En aras de la precisión, la repetiré íntegra; pero conviene tener presente que Gödel no emplea el cálculo para derivar teoremas. Por lo tanto, el lector no necesita dominarlo. *P* es un cálculo predicativo análogo a los CP= del Apéndice IX, pero ajustado a la teoría simple de los tipos (cf. Capítulo 2.4). Los signos primitivos son los conectivos $\sim$ y $\boldsymbol{\vee}$ (negación y disyunción),[10] el signo de cuantificación universal $\Pi$, la constante individual 0 ('cero'), el functor unario $f$ ('el siguiente de'), los paréntesis ) y (, e infinitas variables de distintos tipos, uno para cada entero positivo. Las *variables de tipo 1*, $x_1$, $y_1$, $z_1, \ldots,$ representan números naturales indeterminados; las *de tipo 2*, $x_2$, $y_2$, $z_2, \ldots,$ representan clases de números naturales; las *de tipo 3*, $x_3$, $y_3$, $z_3, \ldots,$ clases de clases de números naturales, etc. Gödel observa que no hace falta admitir variables que representen relaciones, ya que éstas pueden concebirse como clases de clases (por ejemplo, cada relación diádica corresponde a una clase de pares ordenados y cada par ordenado $\langle x,y \rangle$ a la clase de clases $\{\{x\},\{x,y\}\}$). Son *signos de tipo 1* (i) el 0, (ii) cualquier variable de tipo 1, (iii) cualquier signo de tipo 1 precedido por el functor $f$. Un signo de tipo 1 que contiene el 0 —por ejemplo, $ffff0$— se llama *numeral*.[11] Usaré a veces la abrevia-

[10] Le doy al signo de disyunción $\boldsymbol{\vee}$ del cálculo *P* más peso que Gödel para distinguirlo de *nuestro* signo $\vee$, que, como dije en la nota 7, reservo para el discurso metalógico.

[11] Gödel dice 'Zahlzeichen', literalmente 'signo de número'.

tura $f^k0$ para referirme al numeral formado por *k* copias de *f* seguidas de 0. Dicho numeral puede verse como el nombre, en *P,* del número natural *k*. Si $n \geq 2$, un *signo de tipo n* es simplemente una variable de tipo *n*. Si *a* es un signo de tipo *n* ($n \geq 2$) y *b* es un signo de tipo $n$–1, la expresión *a*(*b*) es una *fórmula elemental*. Una *fórmula* de *P* es una fila de símbolos que cae bajo una de las descripciones siguientes:

(1) Una fórmula elemental es una fórmula.
(2) Si *a* es una fórmula, ∼(*a*) es una fórmula (la *negación* de *a*).
(3) Si *a* y *b* son fórmulas, (*a*) ∨ (*b*) es una fórmula (la *disyunción* de *a* y *b*).
(4) Si *a* es una fórmula y *x* es una variable de cualquier tipo, *x*Π(*a*) es una fórmula (una *generalización* de la fórmula *a*); (*a*) es el alcance del cuantificador *x*Π; decimos que la variable *x* está ligada por el cuantificador *x*Π en todas las posiciones que ocupa en la expresión (*a*).

Se advertirá que hay pequeñas diferencias entre *P* y nuestros CP, en cuanto al uso de los paréntesis y la posición relativa del signo de cuantificación y la variable correspondiente. Como aquí estudiaremos, pero no utilizaremos el cálculo *P*, me ha parecido innecesario ajustarlo a la forma acostumbrada. Una variable que no está ligada se dice *libre*. Una fórmula sin variables libres es una *oración*.[12]

El cálculo *P* se organiza como sistema deductivo con infinitos axiomas y dos reglas de inferencia. Las reglas son (i) *modus ponens* y (ii) generalización universal. Gödel las enuncia así: decimos que (i) la fórmula *c* es una *consecuencia inmediata* de las fórmulas *a* y *b* si *a* es la fórmula ∼(*b*) ∨ (*c*) y (ii) *c* es una *consecuencia inmediata* de *a* si *c* es la fórmula *v*Π(*a*), donde *v* es una variable cualquiera. Para facilitar el enunciado de los axiomas, Gödel adopta "las convenciones habituales sobre la omisión de paréntesis" e introduce símbolos auxiliares, utilizados como abreviaturas (en *P*): los conectivos · (conjunción), ⊃ (implicación) y ≡ (equivalencia) y el cuantificador existencial (*Ex*) —que supone definidos en la forma habitual (cf. Apéndice IX.D)— y el signo de identidad = , que define (en la nota 21) así: '$x_n = y_n$'

[12] Gödel dice 'Satzformel', literalmente 'fórmula de oración'.

abrevia a '$x_{n+1}\Pi(\sim x_{n+1}(x_n) \vee x_{n+1}(y_n))$. Los axiomas se clasifican en cinco grupos. Los tres axiomas del grupo I se enuncian así:

I.1 $\sim(fx_1 = 0)$
I.2 $fx_1 = fx_2 \supset x_1 = x_2$
I.3 $x_2(0)\cdot x_1\Pi(x_2(x_1) \supset x_2(fx_1)) \supset x_1\Pi(x_2(x_1))$

Estos axiomas expresan que cero no es el siguiente de ningún número, que el siguiente de un número no es igual al siguiente de ningún otro número y que si cero pertenece a una clase que contiene el siguiente de cualquier número contenido en ella, esa clase contiene todos los números. Corresponden, pues, respectivamente, a los axiomas P4, P3 y P5 de Peano (p. 146).[13] Los restantes axiomas se enuncian mediante esquemas, esto es, reglas para construir determinadas fórmulas de *P*, que serán propiamente los axiomas. Los axiomas formados según los cuatro esquemas del grupo II garantizan la deducibilidad de cualquier tautología (por *modus ponens*). El grupo III comprende todos los axiomas formados reemplazando, en los dos esquemas siguientes, (1) *a* por una fórmula cualquiera $\alpha$, (2) *v* por una variable cualquiera $\xi$, (3) *b* por una fórmula en la que la que $\xi$ no esté libre en ninguna posición y (4) *c* por un signo del mismo tipo que $\xi$, y que no contenga ninguna variable ligada en $\alpha$ por un cuantificador dentro de cuyo alcance figure $\xi$. Los esquemas son:

III.1 $v\Pi(a) \supset a[c/v]$
III.2 $v\Pi(b \vee a) \supset b \vee v\Pi(a)$

Los axiomas del grupo IV desempeñan la función del Axioma de Reducibilidad de Russell y Whitehead, o de los axiomas de comprensión en la teoría de conjuntos. Se obtienen reemplazando en el siguiente esquema (1) *v* por una variable de tipo $n \geq 1$; (2) *u* por una variable de tipo $n$+1 y (3) *a* por una fórmula en que la variable que reemplaza a *u* no ocupa una posición libre:

IV.1 $(Eu)(v\Pi(u(v) \equiv a)$

13 Por su parte, la definición recursiva de numeral —0 es un numeral; si $\eta$ es un numeral, $f\eta$ es un numeral— reemplaza los axiomas P1 y P2.

Por último, el grupo V consta del axioma enunciado a continuación y todas las fórmulas que pueden obtenerse a partir de él mediante el procedimiento que Gödel llama "elevación de tipo".

V.1 $x_1\Pi(x_2(x_1) \equiv y_2(x_1)) \supset x_2 = y_2$

Elevar el tipo de una fórmula dada $\alpha$ consiste, simplemente, en reemplazar las variables $\xi_1, \ldots, \xi_k$, de tipos $t_1, \ldots, t_k$, contenidas en $\alpha$, por variables de $\zeta_1, \ldots, \zeta_k$, de tipos $t_1+n, \ldots, t_k+n$ $(n \geq 1)$.

2º El paso siguiente consiste en asignarle un número de identidad exclusivo a cada signo, fila de signos y lista de filas de signos de *P*. Si $\varphi$ es uno de estos objetos, llamaré el *gödel* de $\varphi$ (abreviado, #[$\varphi$]) al número asignado a $\varphi$ en virtud del procedimiento que describiré enseguida.[14] Doy por sabida la siguiente propiedad de los números naturales: todo número mayor que 1 se puede analizar de una y sólo una manera como producto de factores primos; por ejemplo, $766.360 = 2^3 \times 5^1 \times 7^2 \times 17^1 \times 23^1$. La expresión a la derecha del signo = es la *factorización prima* de 766.360 y los números 2, 5, 7, 17 y 23 son sus *factores primos*. En lo sucesivo, llamaré $p_n$ al *n*-ésimo número primo ($p_1 = 2$, $p_2 = 3$, $p_3 = 5, \ldots$).

Siguiendo a Gödel, estipulamos que #[0] = 1, #[$f$] = 3, #[$\sim$] = 5, #[$\vee$] = 7, #[$\Pi$] = 9, #[(] = 11 y #[)] = 13. Asimismo, asignamos el *n*-ésimo primo mayor que 13 —esto es, $p_{n+6}$— a la *n*-ésima variable de tipo 1. Si $v_t^n$ designa la *n*-ésima variable de tipo *t*, #[$v_t^n$] = $(p_{n+6})^t$ (el *n*-ésimo primo mayor que 13 elevado a la *t*-ésima potencia). Hemos fijado así el gödel de cada signo de *P* de tal modo que dado un número cualquiera *q* se puede determinar mediante un número finito de operaciones aritméticas simples si *q* es el gödel de un signo de *P*: tiene que ser un impar menor que 14 o una potencia de un

[14] La denominación estándar de #[$\varphi$] es 'el *número de Gödel* de $\varphi$'. Esta expresión traduce literalmente las que se usan en alemán y en inglés, idiomas en que el genitivo 'de Gödel' se expresa anteponiendo 'Gödel' a 'Zahl' o 'number' ('número'); pero suena torpe en castellano. En cambio, no me caería mal escribir, casi fonéticamente, 'guédel' (plural, 'guédeles'); pero no me he atrevido a tanto. Por cierto, la denominación 'el gödel de $\varphi$' y el símbolo #[$\varphi$] se aplican también a los números de identidad asignados a los objetos de cualquier cálculo lógico mediante un procedimiento análogo al que se explicará arriba. En el Apéndice XVII, explico el método mucho más simple propuesto por Smullyan (1992) para un cálculo similar a *P*.

primo mayor que 13; además, al averiguarlo, se establece a qué signo corresponde. Supongamos ahora que $\Phi = \alpha_1\alpha_2 \dots \alpha_k$ es una fila de $k$ signos de $P$. Estipulamos que $\#[\Phi] = 2^{\#[\alpha_1]} \times 3^{\#[\alpha_2]} \times \dots \times p_k^{\#[\alpha_k]}$. Por último, si $\mathcal{L} = \langle\Phi_1,\Phi_2, \dots, \Phi_r\rangle$ es una lista de filas de signos de $P$, estipulamos que $\#[\mathcal{L}] = 2^{\#[\Phi_1]} \times 3^{\#[\Phi_2]} \times \dots \times p_r^{\#[\Phi_r]}$. Nuestras estipulaciones garantizan la unicidad del gödel correspondiente a cada fila de signos y a cada lista de filas, así como la posibilidad de recuperar el objeto correspondiente a un número dado con sólo analizarlo en sus factores primos.[15] Por lo tanto, si $n$ es el gödel de un objeto del cálculo, *el objeto con gödel n* es un ente perfectamente determinado, que llamaré †[$n$]. Por otra parte, es claro que, bajo las condiciones prescritas, hay números que no son el gödel de ningún objeto del cálculo $P$.

Sea $K$ una propiedad atribuible a ciertos objetos del cálculo; por ejemplo, la propiedad de ser una variable de tipo 5, o de ser una prueba, o de ser una fórmula no demostrable. Entonces, la clase de números $K^{\#} = \{x{:}x$ es el gödel de un objeto que tiene la propiedad $K\}$ provee una representación numérica de la propiedad $K$. Es natural referirse a los elementos de $K^{\#}$ con un predicado que recuerde al utilizado para designar la propiedad $K$. Gödel (1931) usa con este fin el mismo predicado, impreso en cursivas. Pero aquí, siguiendo la práctica de sus editores (CW, OC), usaré las versalitas. Así, si α es un axioma de $P$, digo que el número #[α] es un AXIOMA. En otras palabras, el número $x$ es un AXIOMA si y sólo si el objeto †[$x$] existe y es un axioma de $P$. Lo mismo vale, *mutatis mutandis*, para las relaciones. Por ejemplo si la fórmula α es una consecuencia inmediata por *modus ponens* de las fórmulas β y γ, digo que #[α] es una CONSECUENCIA INMEDIATA POR *MODUS PONENS* de #[β] y #[γ].

3º Enseguida, Gödel define lo que llama funciones y relaciones *recursivas*. Aunque lo presenta como “una digresión (*eine Zwischenbetrachtung*) que, por de pronto, no tiene nada que ver con el sistema formal $P$” (1931. p. 179), este pasaje —como arriba adelanté— ha tenido repercusiones enormes. Lo

15 Obsérvese, por ejemplo, que un número dado es el gödel de una fila de $k$ signos sólo si es divisible por potencias *impares* de cada uno de los primeros $k$ primos, y es el gödel de una lista de $h$ filas de signos sólo si es divisible por potencias *pares* de cada uno de los primeros $h$ primos.

que Gödel llama aquí "funciones recursivas" es sólo una parte de la familia que hoy conocemos por ese nombre.[16] Pero —como veremos en el Capítulo 2.11— la familia completa se forma con sólo añadir dos sencillas operaciones generadoras a las admitidas por Gödel.[17] Como dije al comienzo, es justificado pensar que esta familia comprende todas las funciones numéricas que es posible computar mediante un algoritmo. Gödel mismo no creyó en un principio que la clase de las funciones computables mediante un algoritmo se pudiera caracterizar así —sólo se convenció después de leer a Turing (1937)— y, por cierto, nunca pretendió que toda función computable fuera recursiva en el sentido restringido definido por él en 1931. Pero cada función recursiva en este sentido restringido es computable mediante un algoritmo que se especifica en su misma definición.

Una función recursiva es una función numérica, esto es, una función cuyo dominio y alcance están contenidos en el conjunto $\mathbb{N}$ de los números naturales. En particular, las funciones recursivas en el sentido restringido de Gödel están definidas en todo $\mathbb{N}$. Esta clase de funciones se distingue sólo por un pequeño detalle de lo que hoy se llama la clase de las *funciones recursivas primitivas*, que definimos así:

---

16 Jesús Mosterín, en su excelente traducción castellana de Gödel, se toma la libertad de escribir "recursiva primitiva" donde el original dice "rekursiv" a secas (OC, p. 62). Esta iniciativa viola la integridad del documento en un punto esencial para el historiador de las ideas. Es verdad que la ilustre matemática húngara Rósza Péter (1934) llamó 'primitiv rekursiv' a las funciones "recursivas" de Gödel 1931, pero ese nombre se usa ahora en un sentido un tanto más amplio: lo que hoy llamamos *funciones recursivas primitivas* coincide con lo que Gödel llama "funciones recursivas" en 1934 (CW I, p. 347; OC, p. 169), mas no con lo que Gödel llamó de ese modo en 1931 (la definición de 1931 omite la cláusula PR3 de la definición de funciones recursivas primitivas que doy enseguida en el texto (*vide* p. 332).

17 Una es la indicada en la referida cláusula PR3: La *proyección* que asigna a cada $r$-tuplo de números naturales su $k$-ésimo término ($1 \leq k \leq r$) es una función recursiva primitiva. Me referiré a la otra en el Capítulo 2.11; pero aquí va un anticipo. Sea $\mu y\Phi(y)$ el más pequeño número natural $y$ que cumple la condición $\Phi(y)$, o el número 0 si ningún número la cumple. Si $\varphi$ es una función recursiva $(n+1)$-ádica, entonces también es recursiva la función $n$-ádica $\psi$ definida por:

$$\psi(x_1, \ldots, x_n) = \mu y(\phi(x_1, \ldots, x_n, y) = 0 \wedge \forall z(z < y \rightarrow \phi(x_1, \ldots, x_n, z) > 0))$$

Kleene (1936) demostró que cualquier función recursiva (en el sentido actual) puede expresarse como una función compuesta de no más de *dos* funciones recursivas primitivas y *una* aplicación de la función $\mu$.

PR1 La función *sucesor* $x \mapsto x + 1$ es recursiva primitiva.

PR2 Si $r$ es un entero positivo, la función *cero* $\langle x_1, \ldots, x_r \rangle \mapsto 0$ es recursiva primitiva.

PR3 Si $r$ es un entero positivo, la función $\langle x_1, \ldots, x_r \rangle \mapsto x_k$ $(1 \leq k \leq r)$ es recursiva primitiva. Esta función se llama la $k$-ésima *proyección* de $\mathbb{N}^r$, simbolizada $\pi^r_k$ o simplemente $\pi_k$.[18]

PR4 Si $h$ es una función recursiva primitiva $m$-aria, y $g_1, \ldots, g_m$ son funciones recursivas primitivas $n$-arias, también es recursiva primitiva la función compuesta:

$$\langle x_1, \ldots, x_n \rangle \mapsto h(g_1(x_1, \ldots, x_n), \ldots, g_m(x_1, \ldots, x_n))$$

PR5a Si $k \in \mathbb{N}$ y $h$ es una función recursiva primitiva 2-aria, también es recursiva primitiva la función 1-aria $f$ definida por la doble condición siguiente:

$$f(0) = k$$
$$f(x+1) = h(x, f(x))$$

PR5b Si $n > 1$, $g$ es una función recursiva primitiva $(n-1)$-aria y $h$ es una función recursiva primitiva $(n+1)$-aria, también es recursiva primitiva la función $n$-aria $f$ definida por la doble condición siguiente:

$$f(0, x_2, \ldots, x_n) = g(x_2, \ldots, x_n)$$
$$f(x_1+1, x_2, \ldots, x_n) = h(x_1, f(x_1, x_2, \ldots, x_n), \ldots, x_n)$$

El lector reconocerá sin dificultad que el esquema de recursión señalado en el texto de Gödel (1931, p. 179) con el número (2) corresponde exactamente a PR5b; el esquema de "sustitución" que él ilustra en la nota 27 (p. 180) corresponde a PR4. Gödel menciona expresamente a PR1 y PR2,[19] y

[18] En rigor, debería indicarse en el nombre de la proyección el valor de $r$, puesto que la $i$-ésima proyección de $m$-tuplos no es la misma función que la $i$-ésima proyección de $n$-tuplos a menos que $n = m$. Pero, en la práctica, este dato está implícito en el contexto y se prefiere no recargar la notación con un signo alusivo a él.

[19] De hecho, Gödel menciona una condición más amplia que PR2, a saber: dado un número cualquiera $c$, la función *constante* $\langle x_1, \ldots, x_r \rangle \mapsto c$ es "recursiva". Pero esta condición se deriva fácilmente de PR1, PR2 y PR4 así: en PR4 tómese $g$ igual a la función cero y $h$ igual a la función sucesor reiterada $c$ veces; entonces la función $\langle x_1, \ldots, x_r \rangle \mapsto c$ es simplemente la función compuesta $\langle x_1, \ldots, x_r \rangle \mapsto h(g(x_1, \ldots, x_r))$.

PR5a es en efecto superflua puesto que se la puede derivar de PR5b y PR2. Gödel omite la condición PR3, la cual, sin duda, amplía la extensión del concepto definido pero no modifica su definición de un modo intuitivamente significativo. En particular, la omisión de PR3 no afecta su demostración de los teoremas que enuncio en el próximo párrafo. Es fácil ver asimismo que las funciones numéricas elementales —suma, multiplicación, exponienciación— son todas recursivas primitivas. Por ejemplo, si ponemos $g(x) = \pi_1(x)$ y $h(x,y,z) = \pi_2(x,y,z) + 1$, la función binaria $f$: $\langle x,y \rangle \mapsto x + y$ satisface la condición PR5b. Como ilustración de PR5a doy siguiente definición de la función factorial $x \mapsto x!$: $0! = 1$, $(n+1)! = (n+1) \cdot n!$.[20]

En adelante, escribiré 'pr-recursiva' por 'recursiva primitiva', y "recursiva" (entre comillas dobles) por 'recursiva en el sentido restringido de Gödel (1931)'. Un relación numérica $n$-ádica $K$ se llama pr-recursiva si su función característica es pr-recursiva; esto es, si es pr-recursiva la función $\chi_K$ definida por:[21]

$$\chi_K(x_1, \ldots, x_n) = 0 \text{ si } \langle x_1, \ldots, x_n \rangle \in K$$

$$\chi_K(x_1, \ldots, x_n) = 1 \text{ si } \langle x_1, \ldots, x_n \rangle \notin K$$

En la próxima nota muestro que estas dos relaciones diádicas son pr-

[20] En este ejemplo, la función definida $f$ es $x \mapsto x!$; la constante $k = 1$ y la función auxiliar $h$ es la multiplicación.

[21] Obsérvese que defino la función característica de $K$ asignando el valor 0 al caso *afirmativo*, el valor 1 al caso *negativo*, tal como hace Gödel. En el Capítulo 1.5 fue oportuno adoptar la convención contraria; cf. p. 43, nota 12.

[22] Para fortalecer su comprensión de lo dicho en el texto, el lector debe probar que las funciones características de las relaciones numéricas 'igual a' y 'menor que' son pr-recursivas. Puede utilizar para eso las cinco funciones pr-recursivas que defino a continuación:

(1) pd $0 = 0$, pd$(x+1) = x$ (pd asigna el predecesor de $x$ a cada número $x > 0$).

(2) $x \dot{-} 0 = x$, $x \dot{-} (y+1) = \text{pd}(x \dot{-} y)$ ($x \dot{-} y$ es igual a $x$ menos $y$ si $x \geq y$; de otro modo, es igual a 0).

(3) $|x - y| = (x \dot{-} y) + (y \dot{-} x)$ ($|x - y|$ es el valor absoluto de la diferencia entre $x$ y $y$).

(4) sg$(0) = 1$, sg$(x+1) = 0$ (sg es la función característica de la clase de números $\{x{:}x > 0\} = \{x{:}x$ es el siguiente de algún número$\}$).

(5) nsg$(0) = 0$, nsg$(x+1) = 1$ (nsg es la función característica de la clase de números $\{x{:}x = 0\} = \{x{:}x$ no es el siguiente de ningún número$\}$).

recursivas: $\{\langle x,y\rangle: x = y\}$ y $\{\langle x,y\rangle: x < y\}$.[22] Una clase pr-recursiva se define del mismo modo, con $n = 1$. En aras de la brevedad, utilizaré la palabra ‘relación’ para designar a clases y relaciones (entiéndase que una “relación 1-ádica” es una clase). Usaré las negritas para representar $n$-tuplos, con $n \geq 1$ arbitrario (Gödel usa con este propósito las letras góticas $\mathfrak{x}$, $\mathfrak{y}$, . . . ). En vez de ‘$x \in K$’ escribo $K(x)$. Gödel demuestra los siguientes teoremas: Si $K$ es una relación pr-recursiva, su complemento $K' = \{x:\neg K(x)\}$ es pr-recursivo. Asimismo, si $K$ y $H$ son relaciones pr-recursivas, también lo es su unión $K \cup H$. Si la función $n$-aria $\varphi$ y la función $m$-aria $\psi$ ambas son pr-recursivas, la relación $(n+m)$-ádica $\{\langle x,y\rangle:\phi(x) = \psi(y)\}$ es pr-recursiva. Por último, si la función $n$-aria $\varphi$ y la relación $(m+1)$-ádica $R$ son pr-recursivas, también los son las relaciones $S$ y $T$ y la función $\Psi$ definidas por:

$$\begin{aligned} &S(x,y) \leftrightarrow \exists z(z \leq \phi(x) \wedge R(z,y)) \\ &T(x,y) \leftrightarrow \forall z(z \leq \phi(x) \rightarrow R(z,y)) \qquad\qquad (\text{✣}) \\ &\Psi(x,y) = \mu z(z \leq \phi(x) \wedge R(z,y)) \end{aligned}$$

donde $\mu z\Phi(z)$ denota el más pequeño número $z$ que cumple la condición $\Phi(z)$, o el número 0 si ningún número cumple dicha condición. Obsérvese que la condición impuesta al número $x$ en la definición de $\Psi$ incluye una cota superior o tope bajo el cual ese número tiene que encontrarse. Esto es indispensable para que la función $\Psi$ sea pr-recursiva (cf. la nota 17). Omitiré las demostraciones de estos teoremas, que no presentan mayor dificultad (véase Gödel 1931, pp. 180-181; OC, pp. 63-65).[23]

---

El lector debe comprobar que la función característica de $\{\langle x,y\rangle: x < y\}$ es $\mathrm{sg}(y \dot{-} x)$ y que la función característica de $\{\langle x,y\rangle: x = y\}$ es $\mathrm{nsg}(|x - y|)$.

23 En todo caso, es claro que la función $\Psi$ y los predicados $S$ y $T$ son computables en el sentido intuitivo. Por ejemplo, para calcular el valor de la función $\Psi$ en el $(n+m)$-tuplo $\langle x_1, \ldots ,x_n,y_1, \ldots ,y_m\rangle$ se procede así: 1º Se calcula el valor de la función pr-recursiva $\phi$ en el $n$-tuplo $\langle x_1, \ldots ,x_n\rangle$; digamos que $\phi(x_1, \ldots ,x_n) = a$. 2º Se calcula la función característica de la relación pr-recursiva $R$ en todos los $(m+1)$-tuplos $\langle 0,y_1, \ldots ,y_m\rangle$, $\langle 1,y_1, \ldots ,y_m\rangle, \ldots , \langle a,y_1, \ldots ,y_m\rangle$ hasta encontrar uno en que dicha función tome el valor 0. Si el primer $(m+1)$-tuplo con esta propiedad es $\langle b,y_1, \ldots ,y_m\rangle$, entonces $b$ es el valor buscado de $\Psi$. Pero si la función característica de $R$ toma el valor 1 en todos los $(m+1)$-tuplos de la lista indicada, entonces el valor buscado de $\Psi$ es 0. Para demostrar que $\Psi$ es pr-recursiva basta describir este procedimiento en términos de las condiciones PR1–PR5.

4º Inmediatamente antes de abordar el paso 3º aludí a la representación numérica de propiedades y relaciones de objetos del cálculo *P* mediante clases y relaciones formadas por los respectivos gödel. Ahora bien, algunas de las propiedades y relaciones sintácticas más importantes —como 'ser una fórmula' o 'ser una prueba'— están representadas de este modo por relaciones pr-recursivas. La definición de las funciones características correspondientes provee, pues, un algoritmo para determinar si ciertos objetos del cálculo —identificados por sus gödel— poseen esas propiedades o están en esas relaciones. Ello era de esperarse, puesto que *P* es visiblemente un cálculo efectivo (en el sentido explicado al comienzo del Capítulo 2.7). La verdadera importancia del uso de la representación numérica de categorías sintácticas mediante relaciones pr-recursivas se mostrará en el paso 5º. Veremos entonces que a cada relación pr-recursiva *n*-ádica le corresponde una fórmula de *P* con *n*-variables libres ($n \geq 1$), que se convierte en una oración demostrable si las variables se reemplazan con los nombres de números que tienen esa relación y en una oración refutable si se las reemplaza con los nombres de números que no la tienen. En virtud de ello, el cálculo *P* contiene una teoría deductiva de su propia sintaxis o, más exactamente, de aquellos aspectos de la misma que son representables mediante relaciones pr-recursivas. Así se cierra de un modo preciso y riguroso el círculo reflexivo que produce el primer teorema de incompletud.

Gödel (1931, pp. 182-186) define 45 relaciones y funciones pr-recursivas que representan categorías sintácticas o sirven para definir a otras que las representan. Las que representan categorías sintácticas se designan con el nombre de la categoría representada impreso en VERSALITAS (de acuerdo con la convención que expliqué al final del 2º). Importa tener presente que las definiciones *no se enuncian en el cálculo P*, sino en la lengua (alemán, castellano) en que se conduce el discurso *sobre el cálculo*. Para mayor claridad, Gödel utiliza abreviaturas e ideogramas, introducidos *ad hoc* o tomados del álgebra ordinaria y del cálculo predicativo de Hilbert y Ackermann (estos últimos se sustituyen aquí con los signos homólogos de nuestro CP). Recomiendo enfáticamente al lector que, para su propio consumo, exprese en palabras las 45 definiciones (así como la definición Nº 46 de una clase de números que no es recursiva). A continuación doy unos pocos ejemplos. Los

dos primeros ilustran las relaciones pr-recursivas *S* y *T* (destacadas arriba con el signo ✣), respectivamente.

1. *x* es divisible por *y* —abreviado *x*/*y*— si y sólo si hay un *z* menor o igual que *x*, tal que $x = y \cdot z$.

2. *x* es primo —abreviado Prim(*x*)— si y sólo si no hay un *z* mayor que 1 y menor que *x* tal que *x* es divisible por *z*.

Los cinco ejemplos siguientes son funciones pr-recursivas análogas a la función Ψ (destacada con ✣).

3. El *n*-ésimo factor primo del número *x* se designa con *n Pr x*.[24] El siguiente esquema de recursión muestra que ésta es una función pr-recursiva:

$$0 \; Pr \; x = 0$$

$$(n+1) \; Pr \; x = \mu y(y \leq x \wedge \text{Prim}(y) \wedge x/y \wedge y > n \; Pr \; x)$$

En palabras: El 0-ésimo factor primo de *x* es 0; el (*n*+1)-ésimo primo contenido en *x* es ya sea (i) el más pequeño número *y* menor o igual que *x* tal que *y* es primo, *x* es divisible por *y* y *y* es mayor que el *n*-ésimo primo contenido en *x*, ya sea (ii) 0, si no existe un número *y* que cumpla estas condiciones.

5. El *n*-ésimo primo es el valor asignado al número *n* por la función pr-recursiva $x \mapsto p_x$.[25] Esta función se define así:

$$p_0 = 0$$
$$p_{n+1} = \mu y(y \leq 1 + p_n! \wedge \text{Prim}(y) \wedge y > p_n)$$

En palabras: El 0-ésimo primo *x* es 0 y el (*n*+1)-ésimo primo es el más pequeño número *y*, menor o igual que el siguiente de $1 \cdot 2 \cdot 3 \cdot \ldots \cdot p_n$, tal que

[24] Uso redonda para la abreviatura de los predicados recursivos (como Prim), y cursiva para la de las funciones recursivas (como *Pr*). Gödel sigue esta norma en las primeras definiciones, pero luego se aparta de ella (vgr. en las definiciones 13 y 17). Ello no tendría la menor importancia si no fuese por su cuestionable práctica de utilizar el mismo símbolo '≡' como signo de equivalencia en la definición de los predicados y como signo de identidad en la definición de las funciones. Por mi parte, usaré '↔' para significar equivalencia y '=' para significar identidad.

[25] En vez de $p_x$, Gödel escribe Pr *x*. Esta abreviatura difiere de la usada por él mismo (y por nosotros) cuando asigna números de identidad a los objetos del cálculo. Además, se parece mucho al nombre de la función binaria definida bajo el número 3.

*y* es un número primo mayor que el *n*-ésimo primo. Para fijar el tope bajo el cual ha de buscarse el (*n*+1)-ésimo primo Gödel aprovecha elegantemente la clásica prueba de que tal número existe, para todo *n* (Euclides, IX.20).

6. La función binaria $\langle n,x \rangle \mapsto n\ Gl\ x$ asigna al par $\langle n,x \rangle$ el exponente del *n*-ésimo factor primo de *x* en la factorización prima de este número (siempre, claro está, que *n* sea mayor que 0 y menor o igual que el número de factores primos de *x*). Se define así:

$$n\ Gl\ x = \mu y(y \leq x \wedge x/(n\ Pr\ x)^y \wedge \neg x/(n\ Pr\ x)^{y+1})$$

En palabras: el número en cuestión es ya sea (i) el más pequeño número $y \leq x$ tal que *x* es divisible por la *y*-ésima potencia del *n*-ésimo factor primo de *x*, pero no es divisible por la (*y*+1)-ésima potencia de ese primo; ya sea (ii) 0 si ese número no existe. La utilidad de esta función reside en que, si *x* es el gödel de una fila de signos o una lista de filas de signos de *P*, *n Gl x* es el gödel del *n*-ésimo signo de esa fila o de la *n*-ésima fila de esa lista.[26]

7. La función $x \mapsto \ell(x)$ se define así:

$$\ell(x) = \mu y(y \leq x \wedge (y\ Pr\ x) > 0 \wedge ((y+1)\ Pr\ x) = 0)$$

Según esto, $\ell(x)$ es el más pequeño número *y* tal que existe un *y*-ésimo mas no un (*y*+1)-ésimo factor primo de *x* (claro está que si *x* no tiene ningún factor primo —esto es, si *x* = 0 ó *x* = 1— $\ell(x) = 0$, de acuerdo con la definición del operador μ). Así, $\ell(x)$ asigna a cada número *x* el número total de sus factores primos. Por lo tanto, si *x* es el gödel de una fila de signos o de

[26] Conviene observar que la explicación en palabras de la definición 6 ofrecida por Gödel es incorrecta. Dice que *n Gl x* es "el *n*-ésimo término (*Glied*) de la secuencia de números correspondiente al número *x*" (1931, p. 182). Alude aquí a la biyección que estableció entre los números y los *k*-tuplos de números (para todo entero positivo *k*) en el paso 2º (y que omití, por superflua, en mi explicación de ese paso). En virtud de ella, cada *k*-tuplo corresponde a un número cuyo factor primo más grande es $p_k$; en particular el *k*-tuplo $\langle n(1),n(2), \ldots ,n(k) \rangle$ corresponde al número $2^{n(1)} \cdot 3^{n(2)} \cdot \ldots \cdot p_k^{n(k)}$. Así, el número 766.360, cuya factorización prima usé arriba como ejemplo, es igual a $2^3 \times 3^0 \times 5^1 \times 7^2 \times 11^0 \times 13^0 \times 17^1 \times 19^0 \times 23^1$ y corresponde, por lo tanto, al 9-tuplo $\langle 3,0,1,2,0,0,1,0,1 \rangle$). Ahora bien, el quinto término de esta secuencia de números es 0, pero según la *fórmula* de la definición 6 —como el lector verificará fácilmente— 5 *Gl* 766.360 = 1 (puesto que 5 *Pr* 766.360 = 23 y 766.360 es divisible por $23^1$ mas no por $23^2$). El error de Gödel es inocuo, porque no hay ningún objeto α del cálculo tal que #[α] = 0.

una lista de filas de signos, $\ell(x)$ es el número total de signos en esa fila o de filas en esa lista; en otras palabras, su "longitud".

8. La función binaria $\langle x,y \rangle \mapsto x * y$ está dada por:

$$x * y = \mu z(z \leq p_{\ell(x)+\ell(y)}{}^{x+\ell(y)y} \wedge \forall n(n \leq \ell(x) \rightarrow n\ Gl\ z = n\ Gl\ x) \\ \wedge \forall n(0 < n \leq \ell(y) \rightarrow (n+\ell(x))\ Gl\ z = n\ Gl\ y))$$

Esta definición[27] está pensada para asegurar (i) que si $x = \#[\xi]$ y $y = \#[\eta]$, donde $\xi$ y $\eta$ son filas de signos de $P$, $x * y = \#[\xi\smile\eta]$, donde $\xi\smile\eta$ es la fila que se forma prefijando $\xi$ a $\eta$, y (ii) que si $x = \#[\mathscr{X}]$ y $y = \#[\mathscr{Y}]$, donde $\mathscr{X}$ y $\mathscr{Y}$ son listas de filas de signos de $P$, $x * y$ es el gödel de la lista que se forma añadiendo la lista $\mathscr{Y}$ a continuación de la lista $\mathscr{X}$. Consideraré el caso (i), dejando el (ii) como ejercicio al lector. Sean pues $\xi$ y $\eta$ dos filas de signos, tales que $x = \#[\xi]$ y $y = \#[\eta]$. Entonces, la fila $\xi\smile\eta$ contiene precisamente $\ell(x) + \ell(y)$ signos y el factor primo más grande de $\#[\xi\smile\eta]$ es $p_{\ell(x)+\ell(y)}$. Verifiquemos que $\#[\xi\smile\eta]$ es el más pequeño número $z$ que cumple las tres condiciones enunciadas en el alcance de $\mu z$. Tenemos que

$$\#[\xi\smile\eta] = p_1{}^{1\,Gl\,x} \cdot \ldots \cdot p_{l(x)}{}^{\ell(x)\,Gl\,x} \cdot p_{\ell(x)+1}{}^{1\,Gl\,y} \cdot \ldots \cdot p_{\ell(x)+\ell(y)}{}^{\ell(y)\,Gl\,y}$$

Este número cumple evidentemente la segunda condición: los exponentes de sus primeros $\ell(x)$ factores primos coinciden uno por uno con los exponentes de los factores primos de $\#[\xi] = x$ (ordenados de menor a mayor). También cumple la tercera: sus últimos $\ell(y)$ factores primos coinciden uno por uno con los exponentes de los factores primos de $\#[\eta] = y$ (ordenados de menor a mayor). Además, este número no puede ser mayor que $p_{\ell(x)+\ell(y)}{}^{x+\ell(y)y}$. En efecto, $\#[\xi\smile\eta] = x \cdot p_{\ell(x)+1}{}^{1\,Gl\,y} \cdot \ldots \cdot p_{\ell(x)+\ell(y)}{}^{\ell(y)\,Gl\,y}$. Obviamente $x \leq p_{\ell(x)+\ell(y)}{}^{x}$. Por consiguiente, bastará mostrar que $p_{\ell(x)+1}{}^{1\,Gl\,y} \cdot \ldots \cdot p_{\ell(x)+\ell(y)}{}^{\ell(y)\,Gl\,y} \leq p_{\ell(x)+\ell(y)}{}^{\ell(y)y}$. Ahora bien, cada uno de los factores primos del número de la izquierda es menor o igual que $p_{\ell(x)+\ell(y)}{}^{y}$. Como el número de esos factores es $\ell(y)$, el producto de todos ellos, elevado al respectivo exponente es menor o igual que $p_{\ell(x)+\ell(y)}{}^{\ell(y)y}$. Por lo tanto, $\#[\xi\smile\eta]$ cumple nuestras tres condicio-

[27] He introducido un cambio inocuo en la definición de Gödel para facilitar la explicación que sigue. El tope para el número $z$ en la definición original es $p_{\ell(x)+\ell(y)}{}^{x+y}$. Al adoptar un tope mayor me será más fácil probar que cierto número que vamos a considerar no lo excede.

nes. Supongamos ahora que hay otro número $u$ que también las cumple. En tal caso, obviamente, $\ell(u) = \ell(x) + \ell(y)$, y para todo $n$, $n\ Gl\ u = n\ Gl\ \#[\xi\smile\eta]$. Si $\#[\xi\smile\eta] \neq u$, por lo menos uno de los factores primos de $\#[\xi\smile\eta]$ no es un divisor de $u$. Por lo tanto, hay un $q \geq 1$ tal que para todo $r$ mayor o igual que $q$ y menor o igual que $\ell(x) + \ell(y)$, $p_r < r\ Pr\ u$ (el $r$-ésimo factor primo de $u$). Como $p_r$ es el $r$-ésimo factor primo de $\#[\xi\smile\eta]$ y su exponente en la factorización prima de $\#[\xi\smile\eta]$ es idéntico al exponente de $r\ Pr\ u$ en la factorización prima de $u$, es claro que $\#[\xi\smile\eta] \neq u$ implica que $\#[\xi\smile\eta] < u$. Por lo tanto, $\#[\xi\smile\eta]$ es el número más pequeño que cumple las condiciones en cuestión.

Después de esta justificación detallada de la definición 8, el lector podrá, espero, entender el empleo de la función $x * y$ en las definiciones siguientes. Componiéndola con la función pr-recursiva $R{:}x \mapsto 2^x$, es posible definir clases pr-recursivas que representan numéricamente distintas categorías de expresiones del cálculo. Obviamente, si $x$ es el gödel de un signo del cálculo, $R(x) = 2^x$ es el gödel de la fila que consta solamente de ese signo. Por lo tanto, si $\alpha$ es una fila de signos y $x = \#[\alpha]$, $R(11) * x * R(13)$ —abreviado $E(x)$— es el gödel de la fila $(\alpha)$; $R(5) * E(x)$ —abreviado $Neg(x)$— es el gödel de la fila $\sim(\alpha)$; si $\zeta$ es una variable y $z = \#[\zeta]$, $R(z) * R(9) * E(x)$ —abreviado $z\ Gen\ x$— es el gödel de la fila $\zeta\Pi(\alpha)$, y así sucesivamente. Combinando varias de estas definiciones, Gödel logra definir, entre otras, la propiedad pr-recursiva Form($x$) —$x$ es una FÓRMULA— que caracteriza a los gödel de las fórmulas de $P$ (Def. 23).

Me detendré un momento a considerar la definición de NUMERAL, el atributo distintivo de los gödel de las expresiones 0, $f$0, $ff$0, $fff$0, . . . , $f^n$0, . . . que representan en el cálculo $P$ a los números naturales 0, 1, 2, 3, . . . , $n$, . . . Primero se define la función pr-recursiva binaria $\langle x,y\rangle \mapsto y\ N\ x$:

16. $0\ N\ x = x$; $(n+1)\ N\ x = R(3) * n\ N\ x$.

Obviamente, si $x$ es el gödel de la fila de signos $\alpha$, $n\ N\ x$ es el gödel de la fila formada por $\alpha$ precedida de $n$ copias del signo $f$. Como $1 = \#[0]$, $n\ N\ R(1)$ es el gödel del numeral que representa el número $n$ en $P$. Por lo tanto, $n\ N\ R(1)$ es el NUMERAL de $n$. En vez de $n\ N\ R(1)$, escribimos con Gödel Z($n$) (Z por *Zahl*, 'número'). Por ejemplo,

$$Z(4) = \#[ffff0] = 2^3 \cdot 3^3 \cdot 5^3 \cdot 7^3 \cdot 11^1 = 101.871.000$$

Recomiendo especialmente al lector que estudie la definición 31, pues la función $Sb(x^{v}_{y})$ a que se refiere tiene un papel destacado en la demostración del Teorema VI. Si $x$, $v$ e $y$ son, respectivamente, los gödel de una fórmula $\phi$, una variable $\upsilon$, y una fila de símbolos $\eta$, $Sb(x^{v}_{y})$ es el gödel de la fórmula que se obtiene al reemplazar $\upsilon$ por $\eta$ en cada una de las posiciones donde $\upsilon$ está libre en $\phi$.

La cadena de definiciones culmina con éstas:

43. $\mathrm{Fl}(x,y,z) \leftrightarrow y = (z \; Imp \; x) \vee \exists u(u \leq x \wedge \mathrm{Var}(u) \wedge (x = u \; Gen \; y)$

Léase: $x$ es una CONSECUENCIA INMEDIATA DE $y$ y $z$ si y sólo si hay entre estos tres números una de las dos relaciones enunciadas a la derecha del signo $\leftrightarrow$. Aquí Var es la clase de los gödel de las variables de $P$ (Def. 12), $u \; Gen \; y$ es la función $R(u) * R(9) * R(y)$ (Def. 15) y $z \; Imp \; x$ es la función $E(R(5) * E(z)) * R(7) * E(x)$ (Defs. 32, 13 y 14). Por lo tanto, si $x$, $y$, $z$ y $u$ son, respectivamente, los gödel de las fórmulas $\xi$, $\eta$, $\zeta$ y de la variable $\upsilon$, $u \; Gen \; y = \#[\upsilon\Pi\eta]$ y $z \; Imp \; x = \#[\sim(\eta)\vee\xi]$. Por lo tanto, $\mathrm{Fl}(x,y,z)$ si y sólo si $x$ es el gödel de una consecuencia inmediata por *modus ponens* de las fórmulas cuyos gödel son $y$ y $z$ o de una consecuencia inmediata por generalización de la fórmula cuyo gödel es $y$ (en cuyo caso $\mathrm{Fl}(x,y,z)$, cualquiera que sea $z$).

44. $\mathrm{Bw}(x) \leftrightarrow (\ell(x) > 0) \wedge \forall n(0 < n \leq \ell(x) \rightarrow \mathrm{Ax}(n \; Gl \; x) \vee \exists u \exists w(0 < u,w < n \wedge \mathrm{Fl}(n \; Gl \; x,u \; Gl \; x,w \; Gl \; x))$

Léase: $x$ es una PRUEBA[28] si y sólo si $x$ tiene al menos un factor primo y el exponente $n \; Gl \; x$ de cada primo $p_n \leq p_{\ell(x)}$ en la factorización prima de $x$ es ya sea (i) el gödel de un axioma de $P$ (véase Def. 42), ya sea (ii) el gödel de una consecuencia inmediata por *modus ponens* de dos fórmulas cuyos gödel son, respectivamente, los exponentes en la factorización prima de $x$ de dos primos $p_u$ y $p_w$ menores que $p_n$, ya sea (iii) el gödel de una consecuencia

[28] En alemán, BEWEISFIGUR, literalmente ‘figura de prueba’. Esta denominación es impropia, puesto que lo que normalmente se llama así, también en alemán, es un *esquema* que, al llenarse con fórmulas del cálculo, genera una prueba. Pero $\mathrm{Bw}(x)$ si y sólo si $x$ es el gödel de una lista de oraciones de $P$ que constituye efectivamente una prueba. Mosterín traduce ‘BEWEISFIGUR’ por ‘DEDUCCIÓN’ (Gödel, OC, p. 70). Esta traducción me parece inobjetable. No la adopto, sin embargo, porque ello me obligaría a usar ‘deducción’ en otros pasajes de este libro como equivalente al inglés ‘proof’ —como, consecuentemente, hace Mosterín en su traducción de Gödel 1934 (OC, p. 182)— y a llamar ‘teoría de la deducción’, en vez de ‘teoría de la prueba’, a la Beweistheorie de Hilbert.

inmediata por generalización de una fórmula cuyo gödel es el exponente en la factorización prima de $x$ de un primo $p_u < p_n$. En otras palabras, Bw($x$) si y sólo si $x$ es el gödel de una prueba en $P$.

45. $x$B$y \leftrightarrow$ Bw($x$) $\wedge$ $y = (\ell(x)\ Gl\ x)$.

Léase: $x$ es una PRUEBA DE $y$ si y sólo si $x$ es una PRUEBA y $y$ es el exponente en la factorización prima de $x$ del máximo divisor primo de $x$. Obviamente, $x$B$y$ si y sólo si hay una prueba cuyo gödel es $x$, e $y$ es el gödel de la última fórmula de esa prueba.

46. Bew($x$) $\leftrightarrow \exists y(y$B$x)$.

Léase: $x$ es una FÓRMULA DEMOSTRABLE si y sólo si hay un número $y$ tal que $y$ es una PRUEBA DE $x$. Obviamente, Bew($x$) si y sólo si existe en $P$ una prueba cuya última fórmula tiene el gödel $x$.

Gödel observa agudamente que —en contraste con todas las definiciones anteriores— no podemos aseverar que la def. 46 defina una clase "recursiva".[29] Obsérvese que, si {$x$:Bew($x$)} fuese "recursiva", podríamos emplear el algoritmo de su función característica para decidir, dada una fórmula $\phi$, si #[$\phi$] pertenece o no a {$x$:Bew($x$)}. Podríamos, entonces, determinar mediante un algoritmo si $\vdash\phi$ o $\nvdash\phi$. En otras palabras, habríamos resuelto el *Entscheidungsproblem* del cálculo $P$. Como veremos en el Capítulo 2.11, en virtud de los hallazgos de Gödel, ni siquiera es posible resolver el *Entscheidungsproblem* del cálculo predicativo de primer orden (que es sólo un fragmento de $P$).

5º El próximo paso consiste en establecer que para cada relación "recursiva"[30] $n$-ádica $R$ y cada $n$-tuplo de números $a$ hay una fórmula de $P$ con $n$ variables libres $\phi(\boldsymbol{\xi})$ tal que, si $\boldsymbol{\alpha}$ es la lista de los numerales correspondientes a $a$, $\vdash\phi(\boldsymbol{\alpha}/\boldsymbol{\xi})$ si $a \in R$ y $\vdash\sim(\phi(\boldsymbol{\alpha}/\boldsymbol{\xi}))$ si $a \notin R$. Como expliqué en la p. 250,

29 El predicado Bew($x$), definido por generalización existencial sobre una de las variables libres del predicado binario recursivo $y$B$x$, es lo que hoy se llama un predicado *recursivamente enumerable*. Esta designación se basa en lo siguiente: puede demostrarse que si un predicado definido de este modo es satisfecho por una clase no vacía de números, dicha clase es el alcance de una función recursiva (general) con dominio $\mathbb{N}$; ésta puede entonces utilizarse para enumerar la clase de números en cuestión (y también, por cierto, cualquier clase de objetos identificados mediante esos números).

30 Recuérdese que llamo así —con el adjetivo entre comillas— a las relaciones recursivas en el sentido restringido de Gödel (1931), o sea, las relaciones cuya función característica es pr-recursiva conforme a las reglas PR1, PR2, PR4, o PR5.

nota 4, una relación numérica con esta propiedad es precisamente lo que Gödel (1931) llama *decidible* (*entscheidungsdefinit*) en *P*. Estableceremos, pues, que las relaciones "recursivas" son *P*-decidibles (como diré para evitar la confusión con otras acepciones de 'decidible' a que me refiero en esa nota). A la luz de lo expuesto en el paso 4º, esto significa que una parte considerable de la sintaxis lógica de *P* es materia de una teoría deductiva formalizable en *P*.

Este resultado está contenido en el Teorema V, el cual dice que para cada relación "recursiva" *n*-ádica *R* hay un número *r* tal que (i) *r* es el gödel de una fórmula con exactamente *n* variables libres †[17], †[19], †[23], . . . , †[$p_{n+6}$];[31] (ii) si *s* es el gödel de la fórmula obtenida al reemplazar la variable †[$p_{k+6}$], en todas las posiciones libres que ésta ocupa en la fórmula †[*r*], por el numeral correspondiente a cierto número $a_k$ ($1 \leq k \leq n$), entonces Bew(*s*) si $R(a_1, \ldots, a_n)$ y Bew(*R*(5) ∗ *E*(*s*)) si $\neg R(a_1, \ldots, a_n)$. Como se recordará, Bew(*s*) si y sólo si la fórmula †[*s*] es demostrable en *P*, y Bew(*R*(5) ∗ *E*(*s*)) si y sólo si la negación de la fórmula †[*s*] es demostrable en *P*.

Como tendremos que referirnos a menudo al Teorema V, conviene disponer de una terminología concisa para aludir a él. Si un número *r* tiene con una relación *R* la conexión arriba descrita, diré que *r* es el número *correspondiente* a *R* y que la fórmula †[*r*] *expresa* a *R* según el Teorema V. Esta última designación (inspirada por Kleene 1952, p. 195) es, por cierto, puramente convencional y arbitraria si *P* es un sistema deductivo inconsistente. Pero si *P* es consistente, la designación se ajusta bien al significado corriente del verbo 'expresar', puesto que, en tal caso, la fórmula obtenida al reemplazar por numerales todas las variables libres de †[*r*] es demostrable en *P* *si y sólo si* los números designados por dichos numerales tienen entre sí la relación *R*.[32]

[31] Recuérdese nuestra convención: si *a* = #[*b*], *b* = †[*a*]. Gödel no especifica cuáles son las *n* variables libres en la fórmula †[*r*]; pero en la nota 38 (1931, p. 186) sugiere el método utilizado arriba.

[32] Consíderese, por ejemplo, una relación diádica *R*(*x*,*y*) y sea $\rho(u_1,v_1)$ la fórmula †[*r*] que la expresa según el Teorema V. Sean *m* y *n* dos números. Según la convención adoptada arriba, los numerales correspondientes son $f^m0$ y $f^n0$. En tal caso, $\neg R(m,n)$ implica que $\vdash\sim\rho(f^m0,f^n0)$ y, por lo tanto, si *P* es consistente, $\vdash\rho(f^m0,f^n0)$ implica que $\neg\vdash\sim\rho(f^m0,f^n0)$ y, por ende, que $\neg\neg R(m,n)$ y que $R(m,n)$.

Aunque el Teorema V propiamente pertenece *a la aritmética* —habla sólo de relaciones entre *números*— el razonamiento que Gödel ofrece como prueba se funda en ciertas propiedades del cálculo *P*. Como observa en una nota, "el Teorema V se basa, naturalmente, en que, si *R* es una relación recursiva, es posible deducir *de los axiomas del sistema P* si un dado *n*-tuplo de números posee o no esa relación" (1931, p. 186, n. 40). Basta probar el Teorema V para toda relación "recursiva" *n*-ádica *R* que satisfaga la condición

$$R(x_1, \dots ,x_n) \leftrightarrow x_1 = \varphi(x_2, \dots ,x_n) \qquad (1)$$

donde $\varphi$ es una función "recursiva"; por cuanto, si *R* es una relación "recursiva" *n*-ádica cualquiera, la relación $R' = \{\langle 0,x_1, \dots ,x_n\rangle : R(x_1, \dots ,x_n)\}$ cumple la condición (1), con $\varphi$ la función característica de *R*. Si $\phi$, en la condición (1), es la función 'el siguiente de' ("recursiva" por PR1), *R* es la relación diádica que subsiste entre dos números si el segundo es el predecesor del primero. En tal caso, la ecuación a la derecha del signo $\leftrightarrow$ en (1) puede obviamente representarse en *P* mediante la fórmula $x_1 = fy_1$. Sea $\alpha$ la fórmula obtenida al reemplazar aquí el par de variables $\langle x_1,y_1\rangle$ por el par de numerales correspondientes a los números $a_1$ y $a_2$. Es claro que $\vdash\alpha$ si $a_1 = a_2+1$ y que $\vdash\sim\alpha$ si $a_1 \neq a_2+1$.[33] Lo mismo vale, *a fortiori*, si $\varphi$, en la condición (1), es una función constante ("recursiva" por PR2).[34] Diré, con Gödel, que una función "recursiva" tiene grado 1 si está definida según PR1 o PR2, y que tiene grado $n$+1 si está definida según PR4 o PR5 a partir de otras funciones "recursivas" entre las cuales hay por lo menos una de grado $n \geq 1$. Supongamos que el Teorema V ha sido demostrado para todas las relaciones "recursivas" que cumplen la condición (1) con una función $\varphi$ de grado menor que un cierto entero positivo *r*. No es difícil ver que, en tal caso, el Teorema V también vale para una relación "recursiva" que cumpla la condi-

[33] Por ejemplo, si $a_1 = 4$ y $a_2 = 3$, $\alpha$ es la fórmula $ffff0 = ffff0$. Si $a_1 = 4$ y $a_2 = 5$. $\alpha$ es la fórmula $ffff0 = ffffff0$. Aquella fórmula y la negación de ésta se deducen fácilmente de los axiomas I.2 y I.1.

[34] Sea, por ejemplo, $\varphi$ la función ($n$-1)-ádica $\langle x_2, \dots ,x_n\rangle \mapsto 2$. Entonces, la ecuación a la derecha del signo $\leftrightarrow$ en (1) puede representarse en *P* mediante la fórmula $x_1 = ff0$. Sea $\alpha$ la fórmula obtenida al reemplazar aquí la variable $x_1$ por el numeral correspondiente a un número cualquiera $a_1$. Es claro entonces que, cualquiera que sea el ($n$-1)-tuplo $\langle a_2, \dots ,a_n\rangle$, $R(a_1, \dots ,a_n) \leftrightarrow a_1 = 2 \leftrightarrow \alpha$ es la fórmula '$ff0 = ff0$'.

ción (1) con una función φ de grado $r$; aunque una demostración rigurosa sería larga y tediosa. Puesto que el Teorema V vale, como vimos, si la función φ es de grado 1, vale, con toda generalidad, para cualquier grado.

6º Nos falta todavía un breve paso antes de abordar por fin el primer teorema de incompletud (Teorema VI). Es claro que si $P$ fuese inconsistente, toda oración sería deducible de los axiomas. Tendríamos entonces que, cualquiera que fuese la oración $\alpha$, $\vdash\alpha$ y $\vdash\sim\alpha$. Por lo tanto, la *consistencia* de $P$ constituye una condición sintáctica *necesaria* para que $P$ sea incompleto (en el sentido de Post). Rosser (1936) demostrará que esta condición también es *suficiente*. Pero Gödel (1931) estableció originalmente la incompletud de $P$ bajo una condición sintáctica más fuerte, que llama "ω-consistencia". Podría pensarse que, en vista del resultado de Rosser, no vale la pena molestarse en estudiarla. Pero —como indico al final del próximo párrafo— la noción de ω-consistencia tiene cierto interés en relación con el programa de Hilbert. Gödel la define de un modo preciso pero un tanto esotérico. Nuestro penúltimo paso consistirá, pues, en explicar esa definición.

Sea $K$ una clase de fórmulas de $P$. La clase D($K$) de las fórmulas deducibles de $K$ es la clase de fórmulas que contiene (i) todas las fórmulas de $K$, (ii) todos los axiomas de $P$ y (iii) toda fórmula de $P$ que se deduzca de dos fórmulas de D($K$) por *modus ponens* o de una fórmula de D($K$) por generalización. (D($K$) es, pues, el conjunto de teoremas de la teoría deductiva que se obtiene añadiendo las fórmulas de $K$ a los axiomas de $P$, o sea, lo que comúnmente se llama la *extensión* de $P$ determinada por $K$). Obsérvese que, según esta definición, la clase de las fórmulas deducibles de los axiomas de $P$ es D(∅). Diré que $K$ es ω-*inconsistente* si y sólo si hay una fórmula $\Phi(u)$, con una sola variable libre $u$, tal que D($K$) contiene a la vez la fórmula $\sim u\Pi(\Phi(u))$ y *todas* las fórmulas que pueden obtenerse reemplazando la variable libre $u$ en $\Phi(u)$ por el nombre de un número natural.[35] Si D(∅) cumple la condición antedicha, *todo* conjunto de fórmulas de $P$ será ω-inconsistente; en tal caso, digo que el propio sistema $P$ es ω-inconsistente. Si la clase de fórmulas $K$ no es ω-inconsistente, diremos que es ω-*consistente*. Diremos, por otra parte, que $K$ es (simplemente) consistente si D($K$) no contiene

[35] Como el lector recordará, la fórmula $\sim u\Pi(\Phi(u))$ "traduce" la fórmula familiar $\neg\forall\xi\Phi(\xi)$ al cálculo $P$.

la negación $\sim\alpha$ de una fórmula $\alpha \in$ D($K$). Obviamente, si $K$ es ω-consistente, también es consistente; pero no es obvio que $K$ sea ω-consistente, si es consistente. Gödel fue, aparentemente, el primero en ver que hay casos de consistencia acompañada de ω-inconsistencia. La representación formal de la aritmética —o de una teoría matemática que la incluya— mediante un sistema deductivo ω-inconsistente sería, por cierto, un desastre. Por lo tanto, para fundamentar la matemática clásica en el espíritu del programa de Hilbert, hay que valerse de medios finitos para certificar no sólo la consistencia, sino también la ω-consistencia de la representación formal adoptada.

Fiel a la práctica que estudiamos en el paso 4º, Gödel define la ω-consistencia como un atributo de ciertas clases de FÓRMULAS, esto es, de las clases de números formados por los gödel de las fórmulas de ciertas clases. El *definiens* se expresa también en términos de clases de números (1931, p. 187). Sea κ la clase de los gödel de nuestra clase arbitraria $K$. Gödel llama Flg(κ) —Flg por *Folgen,* 'consecuencias'— a la *intersección* de todas las clases de números que (i) incluyen la clase κ, (ii) incluyen la clase de los AXIOMAS (esto es, los gödel de los axiomas) y (iii) están cerradas con respecto a la relación CONSECUENCIA INMEDIATA (Def. 43).[36] Obviamente, Flg(κ) es la clase de los gödel de la fórmulas pertenecientes a nuestra clase D($K$). La clase κ de números es ω-consistente según la definición de Gödel si y sólo si la clase $K$ de fórmulas es ω-consistente según la nuestra, esto es, si no existe ningún número $a$ tal que, siendo $a$ el gödel de una fórmula $A(u)$ con la sola variable libre $u$, Flg(κ) contenga a la vez (i) el número $Neg(\#[u]\ Gen\ a) = \#[\sim u\Pi(A(u))]$ y (ii) para todo número $n$, el gödel de la fórmula obtenida al reemplazar $u$ por el numeral $f^n0$ en todas las posiciones libres de $u$ en $A(u)$. (El lector debe persuadirse de que el primer componente de la conjunción utilizada por Gödel en su definición expresa la condición (ii) y el segundo la condición (i); cf. Defs. 13, 15 y 30.).

7º Con los recursos que hemos acumulado en los pasos anteriores, podemos demostrar el Teorema VI, llamado comunmente el primer *Teorema de Incompletud* de Gôdel. En la parte introductoria de su artículo Gödel da una

[36] Decimos que κ está *cerrada* respecto de la relación triádica CONSECUENCIA INMEDIATA si, para cualquier trío de números $x,y,z$ tales que $y,z \in \kappa$ y Fl($x,y,z$), también $x \in \kappa$.

demostración informal del mismo, que reproduje al comienzo de esta sección. Ella ha inspirado la mayoría de las presentaciones posteriores. Como numerosos filósofos han querido reclutar la fuerza de este teorema para diversas causas. me parece importante explicar lo mejor posible la cosa misma, esto es, la demostración formal de Gödel. Se trata de un argumento estrictamente combinatorio, en que partiendo de premisas ya establecidas —casos particulares del Teorema V— se avanza, por pura manipulación de signos, hasta llegar a la conclusión. Hay que cerciorarse de la legitimidad de las sustituciones efectuadas, y para esto hay que tener presente las *correspondencias* entre los objetos designados por los diversos ideogramas que se utilizan. Pero no siempre hace falta *conocer* a los objetos mismos. Hay que atender a las relaciones de implicación o equivalencia entre las oraciones mencionadas, pero no es preciso entender lo que estas oraciones significan. Para refrescar la memoria, doy aquí una lista de ideogramas que se utilizarán con frecuencia:

(i) $\#\alpha$ designa el gödel de $\alpha$ (si $\alpha$ designa un objeto del cálculo $P$);

(ii) $\dagger[x]$ designa el objeto de $P$ cuyo gödel es $x$ (si $x$ es un número que es también un gödel);

(iii) $f^n0$ designa una fila formada por $n$ copias del signo '$f$' seguidas del signo '0'; esta fila es el numeral que representa el número $n$ en $P$;

(iv) $Z(n) = \#f^n0$; en otras palabras, $Z(n)$ es el gödel del numeral que representa al número $n$ en el cálculo $P$;

(v) el 17 y el 19 —los dos primeros primos mayores que 13— son los gödel de dos variables de tipo 1; llamaré $u_1$ a la variable $\dagger[17]$ y $v_1$ a la variable $\dagger[19]$;

(vi) $Sb(y\,{}^{19}_{Z(y)})$ es el gödel de la fórmula que se obtiene al reemplazar la variable $\dagger[19]$ por el numeral de $y$ (esto es, la fila $f^y0$) en todas las posiciones libres que ocupa dicha variable en la fórmula $\dagger[y]$;

(vii) $\dagger[Sb(y\,{}^{19}_{Z(y)})]$ es la fórmula obtenida mediante el reemplazo descrito bajo el (vi);

(viii) $Neg(r)$ es el gödel de la fórmula obtenida prefijando el signo '$\sim$' a la fórmula $\dagger[r]$;

(ix) $17\ Gen\ r$ es el gödel de la fórmula que se obtiene prefijando a la fórmula $\dagger[r]$ la variable $\dagger[17]$ seguida del signo de cuantificación

universal ‘Π’, seguido del paréntesis izquierdo ‘(’ y agregando el paréntesis derecho ‘)’ al final de la fila de signos así obtenida; por lo tanto, †[17 *Gen r*] es la fórmula designada, conforme a nuestras estipulaciones, por el ideograma $u_1\Pi(\dagger[r])$.

El primer Teorema de Incompletud de Gödel (Teorema VI) dice que para cada clase “recursiva” ω-consistente κ de FÓRMULAS hay una FÓRMULA CON UNA VARIABLE LIBRE *r* tal que ni *v Gen r* ni *Neg*(*v Gen r*) pertenecen a Flg(κ) (donde *v* es el gödel de la variable libre en la fórmula †[*r*]). Esto significa que, si *K* es la clase ω-consistente de fórmulas de *P* representada por la clase “recursiva” de números κ, hay una fórmula Φ(*u*), con una sola variable libre *u*, tal que ni *u*Π(Φ(*u*)) ni ~*u*Π(Φ(*u*)) pertenecen a D(*K*). Como ∅ es obviamente “recursiva”, el teorema implica que hay una fórmula Φ(*u*), con una sola variable libre *u*, tal que ni *u*Π(Φ(*u*)) ni ~*u*Π(Φ(*u*)) son deducibles de los axiomas de *P*, de modo que el sistema deductivo *P* es incompleto en el sentido de Post. Pero también implica que esta incompletud *no puede corregirse* ampliando la lista de axiomas, si se requiere que la lista ampliada sea ω-consistente y “recursiva”. Como ya señalé, Rosser (1936) demostró que el requisito de ω-consistencia se puede reemplazar por el simple requisito de consistencia. Además, en lugar de pedir que la lista de axiomas sea “recursiva”, en el sentido restringido de Gödel (1931), basta exigir que sea recursiva, en el sentido más amplio actual, para que el sistema ampliado siga siendo incompleto, si es consistente. Por lo tanto, para que *P* se convierta en un sistema deductivo completo en el sentido de Post habría que añadirle axiomas hasta el punto de que (i) cualquier fórmula de *P* pueda deducirse de ellos o (ii) sea imposible determinar, mediante el algoritmo de una función recursiva, si una fórmula dada es o no es un axioma.

Diré que una lista de fórmulas de *P* es una prueba desde *K* si cada fórmula de la lista es un axioma de *P*, o un miembro de *K*, o una consecuencia inmediata de una o dos fórmulas que la preceden en la lista. Conforme a nuestras definiciones, una fórmula es deducible de *K* si y sólo si es la última fórmula de una prueba desde *K*. La definición de los conceptos $\mathrm{Bw}_\kappa(x)$, $x\mathrm{B}_\kappa y$ y $\mathrm{Bew}_\kappa(x)$ que sigue al enunciado del Teorema VI, adapta a las pruebas desde *K* los conceptos Bw(*x*), *x*B*y* y Bew(*x*) (Defs. 44-46). Si la clase κ de los gödel de las fórmulas de *K* es “recursiva”, también lo son la clase $\{x{:}\mathrm{Bw}_\kappa(x)\}$ de los gödel de las pruebas desde *K* y la clase $\{\langle x,y\rangle{:}x\mathrm{B}_\kappa y\}$ de pares de

números tales que el primero es el gödel de una prueba de la fórmula cuyo gödel es el segundo; mas no la clase $\{x\text{:Bew}_\kappa(x)\} = \text{Flg}(\kappa)$, esto es, la clase de los gödel de las fórmulas deducibles de *K*. Como es obvio, cualquiera que sea la clase de fórmulas *K* representada por la clase de números κ, si la fórmula †[*x*] tiene una prueba, tiene una prueba desde *K*. En otras palabras:

$$\forall x(\text{Bew}(x) \rightarrow \text{Bew}_\kappa(x)) \qquad (2)$$

Gödel llama $Q(x,y)$ a la relación $\neg x\text{B}_\kappa(Sb(y\,{}^{19}_{Z(y)}))$.[37] Esta relación, evidentemente "recursiva", existe entre los números *x* e *y* si *x* no es el gödel de una prueba desde *K* de la fórmula $\dagger[Sb(y\,{}^{19}_{Z(y)})]$ (descrita en la p. 346 bajo el (vii)). Combinando (2) con el Teorema V, concluimos que hay una fórmula con sólo dos variables libres, tal que, si *q* es el gödel de esa fórmula, entonces:

$$\begin{aligned} \neg x\text{B}_\kappa(Sb(y\,{}^{19}_{Z(y)})) &\rightarrow \text{Bew}_\kappa(Sb(q\,{}^{17}_{Z(x)}\,{}^{19}_{Z(p)})) \\ x\text{B}_\kappa(Sb(y\,{}^{19}_{Z(y)})) &\rightarrow \text{Bew}_\kappa(\text{Neg}(Sb(q\,{}^{17}_{Z(x)}\,{}^{19}_{Z(p)}))) \end{aligned} \qquad (3)$$

En lo sucesivo llamo $G(u_1,v_1)$ a la fórmula †[*q*] que expresa según el Teorema V la relación $Q(x,y)$.[38] Sea Ψ la fórmula $\dagger[Sb(y\,{}^{19}_{Z(y)})]$, y sea Θ la

[37] La abreviatura $Q(x,y)$ no vuelve a figurar en la demostración del Teorema VI, pero le sirve más adelante para referirse a la relación $\neg x\text{B}_\kappa(Sb(y\,{}^{19}_{Z(y)}))$ en la demostración del Teorema XI (1931, p. 197). Nosotros también nos valdremos de ella en ese contexto.

[38] Por cierto, '$G(u_1,v_1)$' no *es* la fórmula con gödel *q*, puesto que no es una fórmula del cálculo *P*. '$G(u_1,v_1)$' es el nombre que le doy a la fórmula con gödel *q* en mi discurso sobre el cálculo. Según la práctica descrita en el Apéndice IX.A, combino el nombre '$G(u_1,v_1)$' con signos de *P* para formar nombres que designan otras fórmulas bien determinadas de *P*. Por ejemplo, '$\sim u_1\Pi(G(u_1,v_1))$' designa la fórmula formada por la fila $\sim u_1\Pi$, seguida por el paréntesis izquierdo, seguida por la fórmula llamada $G(u_1,v_1)$, seguida por el paréntesis derecho. Para aligerar la tipografía, llamo $G(u_1,f^p0)$ a la fórmula obtenida al reemplazar la variable $v_1$ por el numeral $f^p0$ en todas las posiciones libres de $v_1$ en $G(u_1,v_1)$. Según la convención descrita al final del Apéndice IX.A, dicha fórmula debiera llamarse $G(u_1,v_1)[f^p0/v_1]$.

fórmula †[$Sb(q\,^{17}_{Z(x)}\,^{19}_{Z(y)})$] obtenida al reemplazar en $G(u_1,v_1)$ las variables libres $u_1$ = †[17] y $v_1$ = †[19] por los numerales $f^x0$ = †[$Z(x)$] y $f^y0$ = †[$Z(y)$], respectivamente. Con esta nomenclatura, las dos implicaciones enunciadas en (3) pueden simbolizarse también así:

$$\begin{aligned} &\neg x\mathrm{B}_\kappa(\#\Psi) \rightarrow \mathrm{Bew}_\kappa(\#\Theta) \\ &x\mathrm{B}_\kappa(\#\Psi) \rightarrow \mathrm{Bew}_\kappa(\mathrm{Neg}(\#\Theta)) \end{aligned} \tag{3*}$$

Según esto, Θ es deducible de $K$ si $x$ no es el gödel de una prueba de Ψ desde $K$ y ∼Θ es deducible de $K$ si $x$ es el gödel de una prueba de Ψ desde $K$.

Sea $p$ = 17 *Gen* $q$ = #[$u_1\Pi(G(u_1,v_1))$] y sea $r = Sb(q\,^{19}_{Z(p)})$. $r$ es el gödel de la fórmula $G(u_1,f^p0)$ que se obtiene al reemplazar en $G(u_1,v_1)$ la variable libre $v_1$ por el numeral $f^p0$ = †[$Z(p)$]. Haciendo las sustituciones pertinentes comprobamos que:

$$Sb(p\,^{19}_{Z(p)}) = Sb((17\ Gen\ q)^{19}_{Z(p)}) = 17\ Gen\ Sb(q\,^{19}_{Z(p)}) = 17\ Gen\ r \tag{4}$$

$$Sb(q\,^{17}_{Z(x)}\,^{19}_{Z(p)}) = Sb(r\,^{17}_{Z(x)}) \tag{5}$$

Si reemplazamos $y$ por $p$ en (3) obtenemos las dos implicaciones siguientes:

$$\neg x\mathrm{B}_\kappa(Sb(p\,^{19}_{Z(p)})) \rightarrow \mathrm{Bew}_\kappa(Sb(q\,^{17}_{Z(x)}\,^{19}_{Z(p)})) \tag{6}$$

$$x\mathrm{B}_k(Sb(p\,^{19}_{Z(p)})) \rightarrow \mathrm{Bew}_k(\mathrm{Neg}(Sb(q\,^{17}_{Z(x)}\,^{19}_{Z(p)}))) \tag{7}$$

Invocando (4) y (5), reformulamos (6) y (7), reemplazando $Sb(p\,^{19}_{Z(p)})$ por 17 *Gen* $r$ y $Sb(q\,^{17}_{Z(x)}\,^{19}_{Z(p)})$ por $Sb(r\,^{17}_{Z(x)})$, respectivamente:

$$\neg x\mathrm{B}_\kappa(17\ Gen\ r) \rightarrow \mathrm{Bew}_\kappa(Sb(r\,^{17}_{Z(x)})) \tag{6*}$$

$$x\mathrm{B}_\kappa(17\ Gen\ r) \rightarrow \mathrm{Bew}_\kappa(Neg(Sb(r\,^{17}_{Z(x)})) \tag{7*}$$

De este resultado inferiremos con Gödel que, si κ es, como hemos supuesto, una clase "recursiva" ω-consistente, ninguno de los números 17 *Gen r* y *Neg*(17 *Gen r*) pertenece a Flg(κ). Así se completa la demostración del Teorema VI. Ahora bien, 17 *Gen r* es el gödel de la fórmula que, según la nomenclatura adoptada, hay que llamar $u_1\Pi(G(u_1,f^p0))$. Así pues, el argumento presentado a continuación demostrará que, si la clase *K* de fórmulas es ω-consistente, ni $u_1\Pi(G(u_1,f^p0))$ ni $\sim u_1\Pi(G(u_1,f^p0))$ son deducibles de *K*. Si $K = \varnothing$, la fórmula llamada $u_1\Pi(G(u_1,f^p0))$ no es demostrable ni refutable en *P*.

Si 17 *Gen r* perteneciera a Flg(κ) tendríamos que hay un número *n* tal que $n\ B_\kappa$ (17 *Gen r*). Por lo tanto, en virtud de (7*), $\mathrm{Bew}_\kappa(Neg(Sb(r^{17}_{Z(n)})))$. Por otra parte, 17 *Gen r* ∈ Flg(κ) significa que hay una prueba desde *K* de la fórmula $u_1\Pi(G(u_1,f^p0))$ y, por ende, también de la fórmula $\dagger[Sb(r^{17}_{Z(n)})]$ obtenida reemplazando en $G(u_1,f^p0)$ la variable libre $u_1$ por $f^n0$. Pero entonces tendríamos que $\mathrm{Bew}_\kappa(Sb(r^{17}_{Z(n)}))$, y κ sería inconsistente y, por ende, ω-inconsistente. Por lo tanto, bajo las hipótesis del Teorema VI, 17 *Gen r* ∉ Flg(κ).

Supongamos, entonces, que *Neg*(17 *Gen r*) ∈ Flg(κ). Como 17 *Gen r* ∉ Flg(κ), es claro que $\forall n\neg(n\ B_\kappa$ (17 *Gen r*)). En virtud de (6*), esto implica que $\forall n\mathrm{Bew}_\kappa(Sb(r^{17}_{Z(n)}))$. Flg(κ) contendría a la vez el número *Neg*(17 *Gen r*) = $\#[\sim u_1\Pi(G(u_1,f^p0))]$ y el gödel de *cada una* de las fórmulas que se obtienen insertando en $G(u_1,f^p0)$ —esto es †[*r*]— los numerales 0, *f*0, *ff*0, . . . en lugar de la variable libre $u_1$. Pero entonces κ sería ω-inconsistente. Por lo tanto, bajo las hipótesis del Teorema VI, *Neg*(17 *Gen r*) ∉ Flg(κ).

Obsérvese que, si la clase κ cumple las condiciones requeridas y le añadimos el número *Neg*(17 *Gen r*), la clase κ′ así formada es a la vez *consistente* y ω-*inconsistente*. κ′ es consistente, puesto que (17 *Gen r*) ∉ Flg(κ) y, por cierto, si κ y por ende ∅ son consistentes,

$$(17\ Gen\ r) \notin \mathrm{Flg}(\{Neg(17\ Gen\ r)\}).$$

κ′ es ω-inconsistente, por la razón aducida para probar que *Neg*(17 *Gen r*) ∉ Flg(κ). Así queda probado, mediante un ejemplo, que la ω-consistencia no equivale a la consistencia.

Gödel subraya que la demostración precedente es *constructiva*, pues todas las aseveraciones existenciales que contiene se basan en el Teorema V, el cual, "como fácilmente se ve, es inobjetable desde un punto de vista intuicionista" (1931, p. 189, n. 45a). Por eso, Gödel recibirá con incredulidad la noticia de que Wittgenstein cuestionaba su descubrimiento. En respuesta a una carta de Karl Menger con citas de Wittgenstein, le escribe el 5 de mayo de 1972:

> En lo que respecta a mi teorema sobre proposiciones indecidibles, los pasajes que usted cita[39] muestran claramente que Wittgenstein no lo entendió (o se hizo como que no lo entendía). Lo interpreta como una especie de paradoja lógica, cuando en realidad es todo lo contrario, a saber, *un teorema matemático perteneciente a una parte absolutamente incontrovertida de las matemáticas: la aritmética finitista o análisis combinatorio*.
>
> (Citado por Wang 1987, p. 49; cursiva mía)

Gödel destaca asímismo que en la demostración del Teorema VI ha invocado solamente dos propiedades del cálculo *P*: (1) la clase de los axiomas y la relación 'consecuencia inmediata' —que epitomiza las reglas de inferencia— se pueden definir recursivamente (paso 4º), si "reemplazamos de algún modo los signos primitivos mediante números naturales" (paso 2º); (2) toda relación "recursiva" es *P*-decidible (paso 5º). "Por lo tanto, en todo sistema formal que satisfaga las condiciones (1) y (2) y sea ω-consistente hay

39 Tomados de Wittgenstein 1956, pp. 50-54 y 176. No viene a cuento examinar aquí estos textos. Baste señalar que las dudas de Wittgenstein responden, en buena parte, al aserto de que la fórmula cuya indemostrabilidad e irrefutabilidad ha sido establecida por Gödel "habla" de sí misma y "dice" que es indemostrable. Este aserto, corriente en las exposiciones de la obra de Gödel, vale, obviamente, para la fórmula $[R_q;q]$ mencionada en el argumento informal presentado en la Sección 2.10.1 (y en el §1 de Gödel 1931), pero no vale para la fórmula †[17 *Gen r*] a que se refiere el Teorema VI. Esta fórmula "hablaría" de sí misma si contuviera el numeral correspondiente a 17 *Gen r*. Pero el único numeral cuya presencia se requiere en †[17 *Gen r*] es $f^p0$, correspondiente al número $p$ = 17 *Gen q* ≠ 17 *Gen r*. Con todo, importando ideas que no pertenecen ni contribuyen en nada al austero argumento combinatorio de Gödel, se puede justificar la "traducción" de la fórmula †[17 *Gen r*] mediante la oración castellana 'la fórmula †[17 *Gen r*] no es demostrable'.

40 En la acepción explicada en la p. 250, nota 4. Véase lo dicho en la p. 342.

proposiciones indecidibles[40] de la forma $\forall xFx$, donde $F$ es una propiedad de números naturales definida recursivamente; y lo mismo ocurre en cualquier extensión de un sistema tal mediante una clase de axiomas recursivamente definible y ω-consistente” (1931, p. 190s.). Gödel anota que tanto el sistema axiomático para la aritmética que consta de los axiomas de Peano, el esquema de definición recursiva PR5 y “las reglas lógicas”, como los sistemas axiomáticos para la teoría de conjuntos de Zermelo-Fraenkel y de von Neumann cumplen las condiciones (1) y (2); y que “la condición (1) está satisfecha por cualquier sistema que tenga las reglas de inferencia habituales y cuyos axiomas se generen (como en *P*) por sustitución a partir de un número finito de esquemas” (1931, p. 191).[41]

El Teorema VI no significa, por cierto, que haya una propiedad de números, descriptible con los medios expresivos de *P* u otro sistema afín, tal que, para cierto número *m* no es verdadero ni falso que *m* la tenga. Tal como ocurre con otros teoremas matemáticos de imposibilidad, todo lo que el Teorema VI quiere decir es que cierta cuestión no puede resolverse *con determinados recursos*. Específicamente, si el sistema considerado es ω-consistente (o siquiera consistente—Rosser 1936), siempre hay una propiedad así cuya posesión por cierto número no se puede demostrar ni refutar *en el sistema*. La imposibilidad comprobada por Gödel es comparable, por ejemplo, a la imposibilidad de construir con regla y compás un segmento recto igual a la circunferencia de un círculo de radio dado (el famoso problema de la “cuadratura del círculo”). Pero habrá de parecernos mucho más grave si creemos que *P* y los sistemas afines comprenden todos los recursos de que dispone el hombre para conocer con certeza una verdad sobre números no incluida ya en la aritmética finitista (Capítulo 2.6).

Completaré esta sección enunciando ciertas consecuencias del Teorema VI que se demuestran en Gödel 1931, §3. Para ello, necesito una definición. Diré con Gödel que una relación (o clase) numérica es *aritmética* si se la puede definir mediante los conceptos de suma, multiplicación e identidad (confinados a los números naturales), utilizando negaciones, disyunciones y

[41] El Apéndice XVI contiene una caracterización general de los sistemas deductivos a los que se aplica una “forma abstracta” del Teorema VI.

generalizaciones (sobre el dominio de los naturales). Una proposición *aritmética* es una aseveración hecha con estos mismos recursos conceptuales.[42] Gödel demuestra sin dificultad que toda relación "recursiva" es aritmética (Teorema VII). Esto implica que para todo problema de la forma $\forall xFx$ —con $F$ recursiva— existe un problema aritmético equivalente (esto es, un problema aritmético cuya solución determina si $\forall xFx$ es verdadera o falsa). Como la prueba del Teorema VII puede formalizarse en $P$, se sigue que cualquiera de los sistemas formales a que se refiere el Teorema VI[43] contiene proposiciones aritméticas indemostrables e irrefutables en el sistema (Teorema VIII). Mediante un argumento interesantísimo que no puedo reproducir aquí, Gödel establece además que todo problema de la forma $\forall xFx$ —con $F$ recursiva— puede reducirse a la cuestión de si una cierta fórmula del cálculo predicativo de primer orden es realizable (Teorema X). En otras palabras, para cada $F$ recursiva hay una fórmula del cálculo predicativo de primer orden que es realizable si y sólo si $\forall xFx$ es verdad. Esto implica que cualquiera de los sistemas formales a que se refiere el Teorema VI contiene fórmulas de primer orden (esto es, fórmulas sin variables ligadas de tipo superior a 1) que no pueden deducirse en el sistema, sin que, por otra parte, sea posible deducir la existencia de un contraejemplo (Teorema IX). Como Gödel (1930) había demostrado que toda fórmula válida del cálculo predicativo de primer orden es deducible en ese cálculo y por ende, *a fortiori*, en $P$ o en una extensión de $P$, las fórmulas a que se refiere el Teorema IX no pueden ser válidas; pero la realizabilidad de su negación es indemostrable en el sistema pertinente ($P$ o una extensión ω-consistente de $P$). Con el Teorema IX Gödel está al borde de establecer que el *Entscheidungsproblem* del cálculo predicativo de primer orden es insoluble (Teorema de Church). Mas para ello haría falta (i) extender a las relaciones recursivas *en general* lo que el Teorema V dice sobre las relaciones "recursivas" y (ii) reconocer que toda función computable mediante un algoritmo es recursiva (Tesis de Church).[44]

[42] Gödel cita dos ejemplos de relaciones aritméticas: La relación diádica '*x es* mayor que *y*' definida por: $x > y \leftrightarrow \forall z\neg(y = x + z)$; y la relación triádica '*x* es congruente con *y* módulo *n*', definida por: $x \equiv y \pmod n \leftrightarrow \exists z(x = y + z{\cdot}n \vee y = x + z{\cdot}n)$.

[43] Vale decir, cualquier sistema deductivo con los signos, fórmulas y reglas de inferencia de $P$ y una clase recursiva ω-consistente de axiomas que incluya los axiomas de $P$.

[44] Conforme a la Tesis de Church, si hubiera un algoritmo para determinar qué oraciones de primer orden son válidas, la clase de esas oraciones y su complemento, la clase de las oraciones de primer orden cuya negación es realizable, serían clases recursivas. Por

### *2.10.3 La indemostrabilidad de la consistencia*

El "segundo teorema de incompletud" (Gödel 1931, Teorema XI) concierne directamente al programa de Hilbert. Simplificando, se lo puede resumir así: si el sistema *P* es consistente, la consistencia de *P* no puede demostrarse en *P*. Por eso, algunos escritores sensacionalistas han concluido que nunca podremos saber si la aritmética es o no contradictoria. Smullyan los ridiculiza observando que "confiar en la consistencia de un sistema sobre la base de que éste puede probar su propia consistencia es tan estúpido como confiar en la veracidad de una persona porque ella dice que nunca miente" (1992, p. 109). Aunque justa, esta observación no disminuye la importancia metodológica del descubrimiento de Gödel. El programa hilbertiano para probar la consistencia de los axiomas de Peano no proponía deducir la consistencia de *P en el mismo P* (o en otro sistema afín que incluya esos axiomas). Se trataba, más bien, de establecer dicha consistencia con recursos esencialmente más débiles. Hasta 1930 se pensaba que todos los recursos necesarios estaban contenidos en la aritmética finitista (Capítulo 2.6). Por otra parte, parecía claro que la aritmética finitista es representable y deducible en *P*. Por lo tanto, si la consistencia de *P* no se podía demostrar en *P*, tampoco se podría establecerla con los recursos de la aritmética finitista.

El Teorema XI —al igual que el Teorema VI— se refiere no solamente a *P* sino a cualquier extensión de *P* mediante una clase recursiva de axiomas. Supondremos que *K* y D(*K*), κ y Flg(κ) son como en la explicación del paso

---

lo tanto, según el Teorema V (extendido), habría en *P* una fórmula con una variable libre $\phi(u)$, tal que $\vdash\phi(n)$ si *n* es el gödel de cualquier oración de primer orden cuya negación es realizable. La prueba de $\phi(n)$ sería una prueba en *P* de que la negación de la oración †[*n*] es realizable. Pero el Teorema IX enseña justamente que hay oraciones de esta clase que no admiten tal prueba. La extensión del Teorema V a las funciones recursivas se sigue inmediatamente de una observación expresada al final de Gödel 1934 (p. 27). Pero durante la visita a Princeton en que dictó las lecciones que forman la base de este trabajo, Gödel no se avino a aceptar la Tesis de Church. Como indica en la "posdata" de 1964 (impresa a continuación de la observación citada), sólo llegó a convencerse a la luz de la obra de Turing (1936). El Teorema de Church sobre la insolubilidad del *Entscheidungsproblem* del cálculo de primer orden —publicado en Church 1936a— es un corolario de otro teorema más general, incluido en Church 1936 (véase el Capítulo 2.11).

6º de la demostración del Teorema VI. *K* es consistente si y sólo si hay una fórmula de *P* que no pertenece a D(*K*). Esto equivale a la existencia de un número *x* tal que *x* es el gödel de una fórmula de *P* y no hay ningún número *y* que sea el gödel de una prueba de esa fórmula desde *K*. Usando las abreviaturas introducidas en la demostración del Teorema VI (cf. Defs. 23 y 45), esta condición necesaria y suficiente para la consistencia de *K* se expresa así: $\exists x \forall y(\mathrm{Form}(x) \wedge \neg(y\mathrm{B}_\kappa x))$. Abreviando aun más, escribo —como Gödel— ‘Wid(κ)’ en lugar de ‘$\exists x \forall y(\mathrm{Form}(x) \wedge \neg(y\mathrm{B}_\kappa x))$’ (no se pierda de vista que ‘Wid(κ)’ es aquí una expresión *castellana* abreviada). Supongamos ahora que κ es una clase “recursiva” de FÓRMULAS. Es claro que, siguiendo una por una las primeras 23 definiciones del paso 4º de la demostración del Teorema VI, es posible construir en *P* una fórmula $F(u_1)$, con una sola variable libre $u_1$, que represente el predicado Form(*x*). Continuando hasta la Def. 45, y combinándola con la definición de κ como clase “recursiva”, se construye una fórmula $B_\kappa(v_1,u_1)$ que representa la relación $(y\mathrm{B}_\kappa x)$. Entonces, la fila

$$\sim u_1\Pi\sim v_1\Pi(F(u_1) \curlywedge \sim B_\kappa(v_1,u_1)) \tag{8}$$

es una oración de *P* que representa el aserto Wid(κ): (8) “dice” en *P* que la clase “recursiva” κ es consistente. Sea *w* el gödel de la oración (8).[45]

El Teorema XI dice que, si κ es una clase recursiva de FÓRMULAS, Wid(κ) $\rightarrow \neg\mathrm{Bew}_\kappa(w)$. Con menos abreviaturas —y referido al cálculo *P* y sus extensiones—, esto quiere decir que, si κ es una clase recursiva que reúne los gödel de todas las fórmulas contenidas en una clase consistente *K*, no existe una prueba desde *K* de la fórmula (8). En particular, si κ = ∅, podemos leer *P* en vez de *K*. En tal caso, la fórmula (8) “dice” que *P* es consistente, y el Teorema XI asevera que esto es así sólo si (8) es indemostrable en *P*.

Por cierto, la oración (8) = †[*w*] no es la única que representa en el cálculo *P* la consistencia de la clase recursiva κ. Supongamos que Wid*(κ) es un enunciado —en castellano— equivalente a Wid(κ) y que *w** es el gödel de una oración que representa a Wid*(κ) en *P*. A menos que se probara que

[45] Gödel (1931, p. 197) escribe: “Sei *w* die SATZFORMEL, durch welche in *P* Wid(κ) ausgedrückt wird”; traducido: “Sea *w* la ORACIÓN mediante la cual Wid(κ) se expresa en *P*”. Conforme a la convención sobre el uso de versalitas, esto quiere decir: “Sea *w* el gödel de la oración mediante la cual Wid(κ) se expresa en *P*”.

$(\sim\dagger[w] \vee \dagger[w^*]) \notin$ D($K$), el Teorema XI, en la forma en que ha sido enunciado no nos fuerza a concluir que $\text{Wid}^*(\kappa) \to \neg\text{Bew}_\kappa(w^*)$. Pero el argumento con que Gödel prueba el Teorema XI no depende específicamente de la estructura de (8). Antes bien, como enseguida veremos, suministra un esquema para probar que $\text{Wid}^*(\kappa) \to \neg\text{Bew}_\kappa(w^*)$, si Wid*(κ) y $w^*$ son lo que dije. Por esta razón, la prueba del Teorema XI destruye toda esperanza de establecer la consistencia de *P* mediante un razonamiento formalizable en *P*, en que dicha consistencia se represente mediante una oración diferente de (8).

El Teorema XI es simplemente una consecuencia del hecho de que el Teorema VI se puede probar formalmente en el cálculo *P*. Construir tal prueba es un trabajo largo y tedioso, que Gödel pospuso para la segunda parte de su artículo (que nunca llegó a escribir). En todo caso, la tarea mayor sería construir las fórmulas de *P* correspondientes a las 45 funciones y relaciones recursivas definidas en el paso 4º de la demostración del Teorema VI, y es claro que esto es pura cuestión de paciencia: la definición semiformal de las respectivas abreviaturas indica la pauta a seguir. Una vez que se dispone de esas fórmulas no es difícil formalizar en *P* el argumento puramente combinatorio que lleva a la conclusión (9). Así, Hilbert y Bernays (1939) probaron rigurosamente en un sistema estrictamente más débil que *P* la oración que representa al Teorema VI en dicho sistema.

Como se recordará, la etapa final de la prueba del Teorema VI consta de dos partes. Se muestra (i) que si κ es consistente, $17\ Gen\ r \notin \text{Flg}(\kappa)$ y (ii) que si κ es ω-consistente, $Neg(17\ Gen\ r) \notin \text{Flg}(\kappa)$. Con las abreviaturas que hemos estado usando, (i) puede escribirse:

$$\text{Wid}(\kappa) \to \forall x\neg(x\text{B}_\kappa(17\ Gen\ r)) \tag{9}$$

Recordemos que $17\ Gen\ r = Sb(p\,{}^{19}_{Z(p)})$, donde $r = Sb(q\,{}^{19}_{Z(p)})$, $p = 17\ Gen\ q$ y $q$ es el número que corresponde según el Teorema V a la relación "recursiva" diádica $\neg x\text{B}_\kappa(Sb(y\,{}^{19}_{Z(y)}))$, abreviada Q($x$,$y$). Reemplazando aquí $y$ por $p$ y $17$ $Gen\ r$ en (9) por $Sb(p\,{}^{19}_{Z(p)})$, comprobamos que (9) también puede escribirse:

$$\text{Wid}(\kappa) \to \forall xQ(x,p) \tag{10}$$

Ahora bien, $Q(x,p)$ es una propiedad "recursiva" expresada en *P* según el Teorema V por la fórmula †[*r*], y †[17 *Gen r*] es precisamente la fórmula de *P* que consta de la variable †[17] seguida del signo Π seguida de la fórmula †[*r*]. Es claro, entonces, que †[17 *Gen r*] "dice" en *P* que $\forall x Q(x,p)$, y que la oración

$$\sim\dagger[w] \vee \dagger[17\ Gen\ r] \tag{11}$$

"dice" en *P* lo que (9) y (10) dicen en castellano. La versión formal de la prueba del Teorema VI en *P* incluirá, pues, una prueba de (11). Por lo tanto, si hubiera una prueba de †[*w*] desde *K*, habría también una prueba de †[17 *Gen r*] desde *K*.[46] Ahora bien, según lo que se vio en la demostración del Teorema VI, tal prueba no puede existir a menos que *K* sea inconsistente. Por lo tanto, si *K* es consistente no puede haber una prueba de †[*w*] desde *K*. Usando abreviaturas: $\text{Wid}(\kappa) \rightarrow \neg\text{Bew}_{\kappa}(w)$. Este es el Teorema XI.

Evidentemente, el enunciado preciso de Wid(κ) —esto es, $\exists x \forall y(\text{Form}(x) \wedge \neg(y\text{B}_{\kappa}x))$— no juega ningún papel en el argumento anterior. Dicho enunciado sirve tan sólo para fijar una fórmula de *P* —a saber, †[*w*]— que represente la consistencia de *P* en *P*. En la prueba formal del Teorema VI en *P* tiene que figurar una fórmula así como antecedente de la implicación (11). Pero, a la luz del argumento con que establecimos la correspondiente implicación (9) en nuestra demostración del Teorema VI en castellano, es claro que —tal como adelanté— †[*w*] podría reemplazarse con cualquier fórmula †[*w**] que represente en *P* un enunciado Wid*(κ) equivalente a Wid(κ).[47]

46 Únase la prueba de (11) —que es una prueba desde ∅ y por ende, *a fortiori*, desde *K*— con la supuesta prueba de †[*w*] desde *K* y úsense estas dos fórmulas como premisas de una inferencia final por *modus ponens*.

47 Como ya he indicado, a los Teoremas VI y XI suele dárseles el nombre de primer y segundo "teorema de incompletud" de Gödel, designándoselos también mediante las abreviaturas G1 y G2, respectivamente. La opinión mayoritaria estima que G2 asestó un golpe de muerte al programa de Hilbert (sección 2.9.4). Ahora bien, G2 depende visiblemente de G1 y no presupone ninguna premisa adicional no admitida por Gödel

en la demostración de G1. En vista de ello, no es descabellado decir que el escollo insalvable en que dicho programa naufraga es G1; sobre todo si se tiene en cuenta que la cuestión de la consistencia de la matemática formalizada fue elegantemente reducida por Hilbert y los suyos al *Entscheidungsproblem*, y G1 establece precisamente que el *Entscheidungsproblem* es insoluble incluso en el caso de la aritmética. Por otra parte, G1 puede extenderse más allá de esos “sistemas afines” a *Principia Mathematica* considerados explícitamente por Gödel, a una clase de teorías formales a las que, por su misma construcción, no se aplica G2. Considérese la propuesta siguiente. Los elementos de la teoría formalizada $T$ se definen recursivamente como de costumbre. Sea $\Pi$ la definición recursiva de *prueba en T* (cf. Apéndice IX.F). Las listas finitas de fórmulas que cumplen con la condición $\Pi$ pueden ordenarse lexicográficamente, o por el tamaño del respectivo gödel. Añadimos a $\Pi$ esta estipulación suplementaria: una lista $\Lambda$ de formulas de $T$ que cumple la condición $\Pi$ es una *prueba en T* sólo si la última fórmula de $\Lambda$ no es la negación de la última fórmula de ninguna prueba anterior a $\Lambda$ en el orden adoptado. Evidentemente, es imposible derivar una contradicción en $T$, y si $T$ es suficientemente rica debiera ser posible derivar en $T$ una oración que represente dicha imposibilidad y de este modo exprese la consistencia de $T$. Michael Detlefsen (1986, 1990) sostiene que, no obstante los resultados de Gödel, el programa de Hilbert podría todavía revivirse empleando tales métodos “atentos a la consistencia” (*consistency-minded*). Reconoce, sí, que “no estamos actualmente en la posición de resolver” la cuestión “muy difícil” de si ello permitirá cumplir dicho programa (1992a, p. 226). Pero pone especial cuidado en refutar a quienes, como Smorynski (1985) vinculan el alegado fracaso del programa de Hilbert directamente con G1, ya que este teorema ciertamente sería aplicable a la referida teoría $T$ si ella es tan rica como sería necesario para que pruebe su propia consistencia.

## 2.11 FUNCIONES COMPUTABLES

Estudiaremos aquí las definiciones precisas de función numérica computable propuestas por Alonzo Church, Alan Turing y otros autores. Según lo convenido en el Capítulo 2.10, llamo *función numérica* a cualquier aplicación $\phi$: $\mathcal{D} \to \mathbb{N}$, tal que $\mathcal{D} \subseteq \mathbb{N}^r$ para algún entero positivo $r$. Es oportuno recordar que cualquier función cuyo valor para un argumento dado pueda calcularse mediante un algoritmo es numérica en este sentido. Por ejemplo, el conocido algoritmo para calcular la raíz cuadrada de un entero positivo $n$ hasta la $q$-ésima decimal lo que en efecto determina es el *numerador* —un entero— de la fracción con denominador $10^q$ que menos dista de $\sqrt{n}$. Y para calcular la raíz cuadrada de un número real no entero tenemos primero que darnos una fracción decimal que lo represente con la aproximación requerida y luego aplicar el algoritmo al numerador de esta fracción.

En el resto de este capítulo diré 'función', en vez de 'función numérica'. Ocasionalmente convendrá distinguir entre funciones *totales*, cuyo dominio $\mathcal{D} = \mathbb{N}^r$ para algún entero positivo $r$, y funciones *parciales*, cuyo dominio $\mathcal{D}$ es una parte propia de $\mathbb{N}^r$. Distinguiré asímismo —al menos, de palabra— las funciones *computables*, que caen bajo los conceptos precisos propuestos por Church, Turing, etc. (ocurre que todos esos conceptos tienen la misma extensión), de las funciones *calculables*, cuyo valor para un argumento dado podría determinarse mediante un algoritmo, conocido o no. Generalmente se admite que toda función computable es calculable en principio, pues los métodos de definición contemplados, aunque pueden ser mucho más complejos que ningún algoritmo tradicional, son siempre combinaciones finitas de procedimientos elementales y cuando alguien nos los explica los reconocemos como algoritmos en el sentido intuitivo ordinario del término.[1] Por

[1] En un artículo reciente, Edward Nelson cuestiona la idea de calculabilidad *en principio* contenida en el "el *dogma finitista* [de] que toda función recursiva es efectivamente calculable" (1993, p. 9). Además de un argumento matemático, dependiente de su *Predicative Arithmetic* (1986), propone la siguiente analogía:

otra parte, la *Tesis de Church* afirma que toda función calculable es computable. Trataremos de esclarecer la naturaleza de este aserto: ¿se trata, como sugiere su nombre, de una conjetura matemática que aún no ha sido demostrada (ni refutada)? ¿o de una convención sobre el uso del lenguaje? ¿o de un enunciado de otra índole?

En las primeras tres secciones de este capítulo consideraré sucesivamente: el concepto general de función recursiva que Gödel propuso en 1934 siguiendo una sugerencia de Herbrand, y su notable simplificación por Kleene (Sección 2.11.1); el concepto de función λ-definible propuesto por Church en 1935 (Church 1936), la Tesis de Church según la cual este concepto coincide con el concepto intuitivo ordinario de función calculable, y el Teorema de Church, que da —bajo este supuesto— una solución negativa al *Entscheidungsproblem* del cálculo predicativo de primer orden (Sección 2.11.2); y la concepción de Turing (1936) de un tipo de "máquinas" que incorporan todos los ingredientes del calcular humano y la definición de computabilidad basada en ella (Sección 2.11.3.). Luego, daré ejemplos y diagramas de máquinas de Turing (Sección 2.11.4), que utilizaré enseguida para demostrar algunos resultados importantes al respecto y, basándome en ellos, el Teorema de Church (Sección 2.11.5).

---

> Los finitistas han verificado [que toda función computable es calculable] en casos sencillos. En otros casos pueden darse por vencidos, pero saben que el cálculo se completaría con sólo persistir lo suficiente. Si les preguntan cuánto es suficiente, su respuesta sería: aproximadamente un número de operaciones igual al numeral que estoy calculando.
>
> Los niños malcriados tienen berrinches. Han comprobado en casos sencillos que esto les procura lo que quieren. En otros casos, pueden darse por vencidos, pero saben que conseguirían lo que quieren con sólo persistir lo suficiente. Si les preguntan cuánto es suficiente, su respuesta sería: hasta que consiga lo que quiero.
>
> Con esto no busco desacreditar a quienes sustentan opiniones diferentes de la mía; simplemente expreso mi opinión de que el finitismo es un sistema autovalidante de creencias que no tiene ninguna base y podría muy bien ser incorrecto.
>
> (Nelson 1993, p. 10)

Nelson no pone en duda la Tesis de Church tal como la enuncio en el texto: toda función efectivamente calculable es computable (recursiva).

### *2.11.1 Funciones recursivas generales*

En 1934 Gödel dictó en el Institute for Advanced Study de Princeton una serie de lecciones sobre "Proposiciones indecidibles en sistemas matemáticos formales". Los apuntes de Kleene y Rosser, repartidos a la sazón a un grupo de subscriptores, luego fueron publicados por Davis (1965) con notas y correcciones de Gödel y una importante "Posdata", fechada el 3 de junio de 1964.[2] En la mayor parte de este escrito Gödel utiliza un concepto de función "recursiva" ligeramente más amplio que en 1931; el nuevo concepto coincide exactamente con lo que en la Sección 2.10.2 llamé *funciones recursivas primitivas* o *pr-recursivas*. Entre las funciones básicas a partir de las cuales toda función "recursiva" puede definirse (aplicando los esquemas PR4 y PR5 de la Sección 2.10.2) se cuentan ahora, no sólo la función sucesor y la función cero (definidas por PR1 y PR2), sino también, para cada entero positivo $r$, la $k$-ésima proyección de $\mathbb{N}^r$ en $\mathbb{N}$ $(1 \leq k \leq r)$, esto es, la función $\langle x_1, \ldots, x_r \rangle \mapsto x_k$ (PR3). Luego, en el §9 y final, Gödel advierte que la función $\phi: \mathbb{N}^2 \to \mathbb{N}$, definida a continuación a partir de dos funciones "recursivas" dadas $\psi$ y $\chi$, en general no es "recursiva" en el sentido especificado por las condiciones PR1-PR5 (vale decir, no es pr-recursiva):

$$\phi(x,0) = \psi(y)$$
$$\phi(x+1,0) = \chi(x)$$
$$\phi(x+1,y+1) = \phi(x,\phi(x+1,y))$$

La definición precedente enseña a calcular la función $\phi$ mediante un procedimiento de recursión simultánea con respecto a dos variables. En una nota añadida en 1964, Gödel recuerda que Ackermann (1928) construyó una función calculable por recursión con respecto a una sola variable y demostró que ella crece más rápido que cualquier función pr-recursiva y, por lo tanto, no puede ser pr-recursiva ella misma.[3]

2 El texto impreso en Davis 1965 figura ahora también en Gödel, CW, I, 346-371, y en traducción castellana en OC, pp. 167-98. Las notas y correcciones con que Gödel suplementó los apuntes en 1934 están incorporadas al texto; las observaciones añadidas en 1964 van entre corchetes.

3 Ackermann (1928) introduce una jerarquía de funciones con valores en $\mathbb{N}$. Una fun-

Gödel se pregunta si no sería posible definir un concepto de función recursiva que abarque todas las formas de recursión imaginables. Enuncia enseguida la definición siguiente, "sugerida por Herbrand en una comunicación privada":

> Si $\phi$ denota una función desconocida y $\psi_1, \ldots, \psi_k$ son funciones conocidas, y las $\psi$ y $\phi$ se toman libremente, del modo más general, las unas como argumento de las otras, y de entre las expresiones resultantes igualamos ciertos pares (*if the* $\psi$*'s and* $\phi$ *are substituted in one another in the most general fashions and certain pairs of the resulting expressions are equated*), entonces, si el sistema resultante de ecuaciones funcionales tiene una y sólo una solución para $\phi$, $\phi$ es una función recursiva.

---

ción es de tipo 1 si sus argumentos son $k$-tuplos de números naturales ($k \geq 1$); de tipo 2, si sus argumentos incluyen funciones de tipo 1; de tipo $n$+1, si sus argumentos incluyen funciones de tipo $n$. Propone el siguiente ejemplo de una función de tipo 2, que utiliza luego para definir la llamada función de Ackermann:

$$\rho_c(f(c),x,0) = x$$
$$\rho_c(f(c),x,\sigma n) = f(\rho_c(f(c),x,n))$$

donde $f$ es cualquier función de tipo 1 y $\sigma n$ designa el siguiente de $n$. Ackermann advierte que el subíndice $c$ indica que $\rho_c(f(c),x,y)$ depende de $f$, $x$ e $y$, pero *no* depende de $c$. Como fácilmente se comprueba, $\rho_c(f(c),a,n)$ es igual al valor de $f$ aplicada iterativamente $n$ veces al argumento $a$. Ackermann introduce además la función auxiliar binaria $\alpha$, definida por $\alpha(x,0) = 0$, $\alpha(x,1) = 1$, $\alpha(x,n+2) = x$. Con estos recursos define la función ternaria $\varphi$ mediante las ecuaciones:

$$\varphi(x,y,0) = x + y$$
$$\varphi(x,y,\sigma n) = \rho_c(\varphi(x,c,n),\alpha(x,n),y)$$

El lector debe comprobar que $\varphi(x,y,1) = xy$ y $\varphi(x,y,2) = x^y$. $\varphi(x,y,3)$ es el resultado de repetir $y$ veces, partiendo del argumento $x$, la operación de elevar al exponente $x$; $\varphi(x,y,4)$ es el resultado de repetir $y$ veces, partiendo del argumento $x$, la operación de elevar al exponente $\varphi(x,y,3)$, etc. La *función de Ackermann* es la función unaria $\psi$ definida, para cada $x \in \mathbb{N}$, por la condición:

$$\psi(x) = \varphi(x,x,x)$$

Ackermann demostró que $\psi(n)$ crece con $n$ más rápidamente que cualquier función de tipo 1. Como todas las funciones pr-recursivas son de tipo 1, $\psi$ no puede ser pr-recursiva. Por otra parte, obviamente, $\psi$ es una función calculable.

Por ejemplo, podríamos tener

$$\begin{aligned}\phi(x,0) &= \psi_1(x)\\ \phi(0,y+1) &= \psi_2(y)\\ \phi(1,y+1) &= \psi_3(y)\\ \phi(x+2,y+1) &= \psi_4(\phi(x,y+2),\phi(x,\phi(x,y+2)))\end{aligned}$$

(Gödel, CW, I, 368; OC, pp. 194s.)

Aunque Gödel no lo diga, debemos dar por supuesto que las funciones pr-recursivas básicas —*sucesor* (PR1), *cero* (PR2) y las *proyecciones* (PR3)— cuentan en todo caso como funciones conocidas (en el ejemplo se utiliza visiblemente la función sucesor, *además de* las $\psi_k$). Gödel adopta la definición anterior con dos importantes restricciones.[4] Requiere (i) que el lado izquierdo de cada ecuación funcional del sistema que determina la función $r$-aria $\phi$ tenga la forma

$$\phi(\psi_{i1}(x_1, \ldots ,x_n),\psi_{i2}(x_1, \ldots ,x_n), \ldots ,\psi_{ir}(x_1, \ldots ,x_n))$$

y (ii) que para cada $r$-tuplo $\langle k_1, \ldots ,k_r\rangle \in \mathbb{N}^r$ haya un y sólo un $m \in \mathbb{N}$ tal que la ecuación $\phi(k_1, \ldots ,k_r) = m$ pueda derivarse del sistema en virtud de las reglas de derivación que Gödel prescribe. La restricción (i) parece destinada sólo a facilitar el enunciado de estas reglas, pues la exigencia de que todas las funciones $\psi_{ij}$ tomen $n$-tuplos como argumentos puede trivialmente cumplirse así: elíjase $n$ de modo que las funciones conocidas que van a uti-

4 Kalmár (1955) construyó una función que satisface la definición atribuida por Gödel a Herbrand, mas no las restricciones añadidas por Gödel. Véase Hermes 1961, §21.7. No hay un algoritmo que permita calcular efectivamente la función de Kalmár. En relación con esto, conviene recordar que, como señalé en la Sección 2.6.3, Herbrand (1931) imponía el requisito siguiente a todas las funciones admisibles en su aritmética formal: "consideradas intuicionistamente" —esto es, "traducidas al lenguaje ordinario, como propiedades de enteros, y no como puros símbolos"— "permiten hacer efectivamente el cálculo" del valor que asignan a cada "sistema particular de números" (esto es, a cada $n$-tuplo, cuando la función considerada es $n$-aria). En vista de esto, me parece que el hallazgo de Kalmár indica que Gödel impuso, con certero instinto, a la definición de función recursiva que Herbrand le había propuesto precisamente las restricciones indispensables para ajustarla al propósito —y a las convicciones— del propio Herbrand.

lizarse para determinar a $\phi$ sean a lo sumo $n$-arias; entonces, si entre ellas hay una función $m$-aria $\psi$ ($m < n$), reemplácesela con la función $n$-aria $\psi'$ definida por $\psi'(x_1, \ldots ,x_m,x_{m+1}, \ldots ,x_n) = \psi(x_1, \ldots ,x_m)$. Las reglas de derivación se definen inductivamente así:

(1*a*) Cualquier expresión obtenida al reemplazar por números naturales todas las variables de una de las ecuaciones del sistema es una ecuación derivada.

(1*b*) Si $\psi_{ij}$ es una de las funciones conocidas que figuran en el sistema, $\langle k_1, \ldots ,k_n\rangle$ es un $n$-tuplo de números naturales y $m$ es un número natural tal que $\psi_{ij}(k_1, \ldots ,k_n)$ es igual a $m$, entonces $\psi_{ij}(k_1, \ldots ,k_n) = m$ es una ecuación derivada.

(2*a*) Si $\psi_{ij}(k_1, \ldots ,k_n) = m$ es una ecuación derivada, entonces la ecuación obtenida al reemplazar $\psi_{ij}(k_1, \ldots ,k_n)$ por $m$ en cualquiera de los lugares en que la expresión $\psi_{ij}(k_1, \ldots ,k_n)$ figura en una ecuación derivada es asimismo una ecuación derivada.

(2*b*) Si $\phi(k_1, \ldots ,k_r) = m$ es una ecuación derivada, donde $k_1, \ldots , k_r$ y $m$ son números naturales, la expresión obtenida al reemplazar $\phi(k_1, \ldots ,k_r)$ por $m$ en una posición ocupada por $\phi(k_1, \ldots ,k_r)$ en el lado derecho de una ecuación derivada es asimismo una ecuación derivada.

Combinada con las reglas de derivación, la restricción (ii) equivale a la condición siguiente: (ii*) los $r$-tuplos $\langle k_1, \ldots ,k_r\rangle$ que son los argumentos posibles de $\phi$ pueden ordenarse de tal modo que para computar el valor de $\phi$ correspondiente a cierto $r$-tuplo $\langle k_1, \ldots ,k_r\rangle$ sea preciso conocer, a lo sumo, los valores de $\phi$ correspondientes a $r$-tuplos que *preceden* a $\langle k_1, \ldots ,k_r\rangle$ en dicho orden.

Llamamos *función recursiva general* a toda función numérica que satisfaga estos criterios de Gödel (1934). En la literatura se ha solido llamarlas funciones recursivas “en el sentido de Herbrand-Gödel”. Adviértase que, en virtud de la restricción (ii), las funciones recursivas generales son funciones *totales*. Esta fue históricamente la primera de las nociones equivalentes de función computable a que aludí arriba. Es obvio que toda función recursiva general es calculable mediante un algoritmo. Gödel se preguntó, a la inversa, si las funciones calculables son todas recursivas, pero no aventuró una

respuesta positiva a esta pregunta (cf. Church 1936, p. 356, n. 18). Años más tarde, en la "Posdata" de 1964 a las lecciones de 1934, Gödel declara que Turing (1936) fue el primero que dio un análisis satisfactorio de 'procedimiento mecánico' ("alias 'algoritmo' o 'procedimiento de cómputo' o 'procedimiento combinatorio finito'"—Gödel, CW, I, 369; OC, p. 196). Basándose en ese análisis, Turing construyó su definición precisa de función computable que tiene, demostrablemente, la misma extensión que el concepto de función recursiva general. En vista de esto, Gödel acabó aceptando que todas las funciones calculables son recursivas en el sentido definido por él en 1934.

Gracias a un hallazgo de Stephen C. Kleene (1936) se puede dar una caracterización de las funciones recursivas generales mucho más perspicua y elegante que la original de Gödel. Kleene demostró que basta suplementar las operaciones PR4 y PR5 con una sencilla operación adicional para generar todas las funciones recursivas generales a partir de las funciones pr-recursivas básicas (PR1-PR3). Se trata de la operación de buscar, partiendo de 0, el número natural más pequeño que cumple una condición pr-recursiva dada (esto es, que posee una propiedad dada cuya función característica sea pr-recursiva). En la p. 334 (✣), vimos que Gödel (1931) utilizó esta operación *restringida a un segmento de* $\mathbb{N}$ *acotado por arriba* para definir un tipo de funciones pr-recursivas que luego figuran en la construcción del predicado pr-recursivo Bw($x$) ('$x$ es el gödel de una prueba'). Gödel demostró que si $\phi$ es una función pr-recursiva $n$-aria y $R$ es una relación pr-recursiva ($m$+1)-aria, también es pr-recursiva la función que asigna a cada ($n$+$m$)-tuplo $\langle x_1, \ldots, x_n, y_1, \ldots, y_m \rangle$ el más pequeño número $x \leq \phi(x_1, \ldots, x_n)$ que cumple la condición $R(y_1, \ldots, y_m, x)$, o el número 0 si no hay un $x \leq \phi(x_1, \ldots, x_n)$ que la cumpla. Con el simbolismo utilizado en la p. 334, el valor de esta función en el argumento $\langle x_1, \ldots, x_n, y_1, \ldots, y_m \rangle$ se llama $\mu x(x \leq \phi(x_1, \ldots, x_n) \wedge R(y_1, \ldots, y_m, x))$. Kleene se interesa por la operación de búsqueda simbolizada con $\mu$, cuando *no está restringida por una cota superior o tope*. Supongamos, pues, que hay por lo menos un número que cumple la condición $R(y_1, \ldots, y_m, x)$ y sea $\mu x(R(y_1, \ldots, y_m, x))$ el número más pequeño que la cumple. Si $\chi_R$ es la función característica de la relación $R$, este número se designa asimismo con la expresión $\mu x(\chi_R(y_1, \ldots, y_m, x) = 0)$. Kleene (1943, p. 45) muestra que si $\chi_R$ es una función recursiva general, la función $\langle y_1, \ldots, y_m \rangle \mapsto \mu x(\chi_R(y_1, \ldots, y_m, x) = 0)$ también es recursiva general. Para ello, introdu-

ce una función auxiliar que llamaré $\psi$, definida por las ecuaciones

$$\psi(0,y_1,\ldots,y_m,x) = x \tag{1}$$
$$\psi(z+1,y_1,\ldots,y_m,x) = \psi(\chi_R(y_1,\ldots,y_m,x+1),y_1,\ldots,y_m,x+1) \tag{2}$$

y luego define la función $\phi:\langle y_1,\ldots,y_m\rangle \mapsto \mu x(\chi_R(y_1,\ldots,y_m,x) = 0)$ mediante la ecuación

$$\phi(y_1,\ldots,y_m) = \psi(\chi_R(y_1,\ldots,y_m,0),y_1,\ldots,y_m,0) \tag{3}$$

Esta ecuación define a $\phi$ en la forma requerida por Gödel, en términos de las dos funciones "conocidas" $\psi$ y $\chi_R$. Para ver que $\phi$ realmente es la función que he dicho, considérese un $m$-tuplo cualquiera $q = \langle q_1,\ldots,q_m\rangle$. Calculemos el valor de $\phi(q)$ mediante reemplazos en la ecuación (3). Primero reemplazamos $\chi_R(q,0)$ por su valor. Si es igual a 0, la ecuación (1) nos da la respuesta buscada: $\phi(q) = \psi(0,q,0) = 0$. Si $\chi_R(q,0) = 1$, tenemos que $\phi(q) = \psi(1,q,0) = \psi(\chi_R(q,1),q,1)$, por la ecuación (2). En la última expresión, reemplazamos $\chi_R(q,1)$ por su valor. Si es igual a 0, la ecuación (1) nos dice que $\phi(q) = 1$. En cambio, si $\chi_R(q,1) = 1$, tenemos que $\phi(q) = \psi(1,q,1) = \psi(\chi_R(q,2),q,2)$, por la ecuación (2). El proceso se repite hasta encontrar un número $p$ tal que $\chi_R(q,p) = 0$. (Como se recordará, hemos supuesto que hay por lo menos un número tal). En virtud de la ecuación (1), es claro que $\phi(q) = p$. Porque los números se consideran consecutivamente, de menor a mayor, a partir de 0, es claro que el primer número $p$ que cumpla la condición $\chi_R(q,p) = 0$ es también el más pequeño que la cumple. De modo que $\phi(q) = p = \mu x(\chi_R(q,x) = 0)$.

Kleene (1936) demostró el siguiente teorema: Si $\varphi$ es una función recursiva general $r$-aria, entonces, para cada $r$-tuplo $\langle x_1,\ldots,x_r\rangle \in \mathbb{N}^r$,

$$\phi(x_1,\ldots,x_r) = \psi(\mu y(\chi(x_1,\ldots,x_r,y) = 0))$$

donde $\psi$ es una función pr-recursiva $r$-aria y $\chi$ es una función pr-recursiva $(r+1)$-aria tal que la ecuación $\chi(x_1,\ldots,x_r,y) = 0$ tiene por lo menos una solución $y \in \mathbb{N}$, para cada $r$-tuplo $\langle x_1,\ldots,x_r\rangle \in \mathbb{N}^r$. La demostración de este teorema no es demasiado difícil, pero toma tiempo, y no puedo darla aquí.[5]

[5] Recomiendo la demostración de Hermes 1961, §21 (hay traducción inglesa).

En virtud del mismo, toda función recursiva general puede expresarse como una función compuesta —en el orden indicado— de dos funciones pr-recursivas y *una* aplicación del operador $\mu$. Como toda función pr-recursiva es evidentemente recursiva general, este hallazgo de Kleene nos permite reemplazar la definición gödeliana de función recursiva general por las seis estipulaciones siguientes:

R1 La función *sucesor* $x \mapsto x + 1$ es recursiva general.

R2 Si $r$ es un entero positivo, la función *cero* $\langle x_1, \ldots, x_r \rangle \mapsto 0$ es recursiva general.

R3 Si $r$ es un entero positivo, la función $\langle x_1, \ldots, x_r \rangle \mapsto x_k$ $(1 \leq k \leq r)$ es recursiva general. Esta función se llama la $k$-ésima *proyección* de $\mathbb{N}^r$, simbolizada $\pi^r_k$ o simplemente $\pi_k$.[6]

R4 Si $h$ es una función recursiva general $m$-aria, y $g_1, \ldots, g_m$ son funciones recursivas generales $n$-arias, también es recursiva general la función compuesta:

$$\langle x_1, \ldots, x_n \rangle \mapsto h(g_1(x_1, \ldots, x_n), \ldots, g_m(x_1, \ldots, x_n))$$

R5 Si $g$ es una función recursiva general $(n-1)$-aria (una constante, si $n = 1$) y $h$ es una función recursiva general $(n+1)$-aria, también es recursiva general la función $n$-aria $f$ definida por la doble condición siguiente:

$$f(0, x_2, \ldots, x_n) = g(x_2, \ldots, x_n)$$
$$f(x_1+1, x_2, \ldots, x_n) = h(x_1, f(x_1, x_2, \ldots, x_n), \ldots, x_n)$$

R6 Si $h$ es una función recursiva general $(r+1)$-aria $(r \geq 1)$ y para cada $r$-tuplo $\langle x_1, \ldots, x_r \rangle \in \mathbb{N}^r$ hay un $y \in \mathbb{N}$ tal que $h(x_1, \ldots, x_r, y) = 0$, también es recursiva general la función $r$-aria $g$ definida por la condición siguiente:

$$g(x_1, \ldots, x_r,) = \mu y(h(x_1, \ldots, x_r, y) = 0)$$

[6] Como indiqué en la p. 332, nota 18, una notación rigurosa debería incluir el valor de $r$ en el símbolo de la proyección. En la práctica, este dato se omite porque normalmente está implícito en el contexto.

La condición R6 exige que la ecuación $h(x_1, \ldots ,x_r,y) = 0$ tenga por lo menos una solución $y$ para cada $r$-tuplo $\langle x_1, \ldots ,x_r\rangle$. Este requisito es indispensable para que la función $g$ cumpla el requisito (ii) prescrito por Gödel en su definición de las funciones recursivas generales, esto es, para que, dado un $r$-tuplo $\langle k_1, \ldots ,k_r\rangle \in \mathbb{N}^r$ haya un y sólo un $m \in \mathbb{N}$ tal que $g(k_1, \ldots ,k_r) = m$. Pero, visiblemente, se trata de una exigencia muy artificial. Si la eliminamos, obtenemos el concepto de *función recursiva parcial*, introducido por Kleene (1938), el cual constituye nuestra noción más amplia de función recursiva. Kleene (1943) lo define así:

> Una función $\phi(x_1, \ldots ,x_n)$ es *recursiva parcial* si existe un sistema E de ecuaciones que la define recursivamente en el sentido que se explica a continuación. Un sistema E de ecuaciones *define recursivamente* una función recursiva parcial $n$-aria si, para cada $n$-tuplo $\langle x_1, \ldots ,x_n\rangle$ de números naturales se puede derivar de E, conforme a las reglas [de derivación], *a lo sumo para un* numeral $\mathsf{x}$ una ecuación de la forma $\mathsf{f}(\mathsf{x}_1, \ldots ,\mathsf{x}_n) = \mathsf{x}$, donde $\mathsf{x}_1, \ldots ,\mathsf{x}_n$ son los numerales que representan a los números naturales $x_1, \ldots ,x_n$. La función $n$-aria definida por E en este caso es la función $\phi$ cuyo valor $\phi(x_1, \ldots ,x_n)$ para el argumento $\langle x_1, \ldots ,x_n\rangle$ es *el número natural x representado por el numeral* $\mathsf{x}$ *si ese numeral existe, y de otro modo está indefinido*.[7]

En virtud de esta definición, toda función recursiva general es recursiva parcial. Por eso, en vez de 'función recursiva parcial', se suele decir simplemente *función recursiva*.

7 Construí la definición arriba transcrita haciendo los reemplazos indicados por Kleene (1943, p. 50s.) en su definición gödeliana de función recursiva general (1943, pp. 44s.). En vez de "las reglas [de derivación]", Kleene dice "las reglas R1 y R2", con lo cual alude, no por cierto a *nuestras* reglas R1 y R2, sino a su propia versión (p. 43) de las reglas de derivación de Gödel.

### *2.11.2 La Tesis y el Teorema de Church*

El primero en reclamar que toda función efectivamente calculable es una función recursiva general fue Alonzo Church (1936). En este trabajo —un resumen de cuyos resultados se publicó en 1935— Church se propone dar "una definición de calculabilidad efectiva que corresponda satisfactoriamente a la noción intuitiva un tanto vaga" (p. 346) que solía invocarse en la literatura matemática. Se vale para ello del concepto de *función λ-definible* —desarrollado por él y Kleene— cuya extensión demostrablemente coincide con la del concepto gödeliano de función recursiva general.

El concepto de función λ-definible se introduce en términos de un sistema formal extraordinariamente simple, que llamaré el *λ-cálculo*. Los símbolos del formalismo son el operador λ, los tres pares de paréntesis (, ), [, ], {, }, y una lista infinita de *variables*, que identifico con $x$, $x_{|}$, $x_{||}$, . . . , aunque, siguiendo a Church, las llamaré $a$, $b$, $c$, . . . Una *fórmula* es cualquier secuencia finita de símbolos. Usaré, como Church, caracteres en negrita como variables metalingüísticas que representan fórmulas. (Por lo tanto, en la presente sección, una letra en negrita *no representa un n-tuplo*). Los conceptos de *fórmula bien formada* (fbf), variable *libre* y variable *ligada* se definen inductivamente como sigue. (i) Si $\mathbf{x}$ es una variable, $\mathbf{x}$ es una fbf y la variable $\mathbf{x}$ está libre en la posición que ocupa en esa fbf. (ii) Si $F$ y $X$ son fbfs, $\{F\}(X)$ es una fbf; si la variable $\mathbf{x}$ está libre (ligada) en cierta posición que ocupa en $F$ o $X$, también está libre (ligada) en la posición correspondiente en $\{F\}(X)$. (iii) si $M$ es una fbf y $\mathbf{x}$ es una variable libre en $M$, $\lambda\mathbf{x}[M]$ es una fbf, $\mathbf{x}$ está ligada en todas las posiciones que ocupa en $\lambda\mathbf{x}[M]$ y cualquier variable $\mathbf{z} \neq \mathbf{x}$ que esté libre (ligada) en $M$ también está libre (ligada) en la posición correspondiente en $\lambda\mathbf{x}[M]$.

Church introduce varias abreviaturas: si $F$ consiste en (o se representa mediante) un solo signo, escribe $F(X)$ en vez de $\{F\}(X)$, $F(X,Y)$ en vez de $\{\{F\}(X)\}(Y)$, etc.; en vez de $\lambda\mathbf{x}_1[\lambda\mathbf{x}_2[\ldots\lambda\mathbf{x}_n[M]\ldots]]$ escribe $\lambda\mathbf{x}_1\mathbf{x}_2\ldots\mathbf{x}_n\cdot M$ o, si $M$ no empieza con una variable, simplemente $\lambda\mathbf{x}_1\mathbf{x}_2\ldots\mathbf{x}_nM$. Luego distingue una secuencia de fbfs construidas según un procedimiento uniforme, a las que designa abreviadamente con números, que indican el respectivo lugar en la secuencia. Según esto,

| | | |
|---|---|---|
| 0 | abrevia a | $\lambda ab\cdot a(b)$ |
| 1 | abrevia a | $\lambda ab\cdot a(a(b))$ |
| 2 | abrevia a | $\lambda ab\cdot a(a(a(b)))$ |

y así sucesivamente.[8] A estas fbfs las llamaré *numerales*. Si $F$, $G$ y $H$ son fórmulas, designo con $\mathsf{S}^{F}_{G}H|$ la fórmula que resulta al reemplazar $F$ por $G$ en todas las posiciones que $F$ ocupa en $H$.

Hay tres reglas para derivar una fbf $F$ de una dada fbf $G$. Si $M$ y $N$ son fórmulas y $\mathbf{x}$ y $\mathbf{z}$ son variables, la fórmula derivada $F$ puede ser:

I La fbf resultante al reemplazar en $G$ una parte $\lambda\mathbf{x}[M]$ por $\lambda\mathbf{z}[\mathsf{S}^{\mathbf{x}}_{\mathbf{z}}M|]$, donde $\mathbf{z}$ es una variable que no figura en $M$.

II La fbf resultante al reemplazar en $G$ una parte $\{\lambda\mathbf{x}[M]\}(N)$ por $\mathsf{S}^{\mathbf{x}}_{N}M|$, siempre que las variables ligadas en $M$ difieran de $\mathbf{x}$ y de todas las variables libres en $N$.

III La fbf resultante al reemplazar en $G$ una parte $\mathsf{S}^{\mathbf{x}}_{N}M|$ (no precedida inmediatamente por $\lambda$) por $\{\lambda\mathbf{x}[M]\}(N)$, siempre que las variables ligadas en $M$ difieran de $\mathbf{x}$ y de todas las variables libres en $N$.

Si $G$ es una fbf y $F = G$, o $F$ se deriva de $G$ según una de estas tres reglas, decimos que $G$ es *inmediatamente convertible* en $F$, simbólicamente $G \Rightarrow F$.[9] $G$ es *convertible* en $F$ —abreviado $G$ *conv* $F$— si $G \Rightarrow F$ o si hay una fbf $H$ tal que $G$ *conv* $H$ y $H \Rightarrow F$. Si $G$ *conv* $F$, hay pues una secuencia finita de fbfs $G \Rightarrow H_1 \Rightarrow H_2 \Rightarrow \ldots \Rightarrow H_n \Rightarrow F$, en la que cada fbf (excepto la última) es inmediatamente convertible en la siguiente en virtud de alguna de las reglas I-III. Tal secuencia es una *conversión*. Llamo *sustitución de variables* a una conversión que sólo incluye aplicaciones de la regla I. Llamo, con Church, *reducción* a una conversión que combina aplicaciones de la regla I con una aplicación única de la regla II (pero no aplica la regla III).

8 Formalmente (sin abreviaturas) 0 es la fbf $\lambda x[\lambda x_{|}[\{x\}(x_{|})]]$, 1 es $\lambda x[\lambda x_{|}[\{x\}(\{x\}(x_{|}))]]$, 2 es $\lambda x[\lambda x_{|}[\{x\}(\{x\}(\{x\}(x_{|})))]]$, etc. Adviértase que en vez de 0, 1 y 2, Church escribe 1, 2 y 3, respectivamente, porque en todo su artículo identifica los números naturales con los enteros positivos (cuenta desde uno, y no como nosotros desde cero). Por la misma razón, Church estipula que la función característica de un conjunto $C$ asigna el valor 2 a los elementos de $C$ (y el valor 1 a los objetos que no pertenecen a $C$).

9 Church no emplea este simbolismo.

Una función $\varphi:\mathbb{N} \to \mathbb{N}$ es *λ-definible* si hay en el λ-cálculo una fbf *F* tal que, si $\varphi(m) = r$, y *m* y *r* son los numerales que, según la convención adoptada, designamos con los números *m* y *r*, respectivamente, entonces, $\{F\}(m)$ *conv r*. Esta definición se extiende sin dificultad a funciones cuyo dominio es una potencia cartesiana de ℕ. Generalmente, $\psi:\mathbb{N}^q \to \mathbb{N}$ es *λ-definible* si hay en el λ-cálculo una fbf *F* tal que, si $\varphi(m_1, \ldots, m_q) = r$, y $m_1, \ldots, m_q, r$ son los numerales que designamos con los números $m_1, \ldots, m_q, r$, respectivamente, entonces, $\{F\}(m_1, \ldots, m_q)$ *conv r*.[10]

Antes de seguir adelante, defino dos conceptos auxiliares que utilizaremos luego. Una fbf está *en forma normal* si no contiene ninguna parte de la forma $\{\lambda\mathbf{j}[M]\}(N)$. Si la fbf *F* está en forma normal, cualquier fbf que sea parte de *F* está en forma normal. Decimos que una fbf *F* tiene la forma normal *G* si *F conv G* y *G* está en forma normal. Una fbf puede tener varias formas normales, pero todas ellas son mutuamente convertibles por sustitución de variables. Church indica un modo de distinguir, entre las varias formas normales de una fbf dada, una *forma normal principal* única, determinada por la selección de las variables que figuran en ella y el orden en que aparecen. La forma normal principal se define de tal modo que todo numeral es una forma normal principal. Church destaca que "en el caso de cualquier función λ-definible de enteros positivos,[11] el proceso de reducción de fórmulas a la forma normal provee un algoritmo para calcular efectivamente los valores particulares de la función" (p. 349).

Después de presentar el λ-cálculo, Church da reglas para asignar un número de identidad —un gödel— a sus símbolos y fórmulas. Luego define lo que llama *funciones recursivas* en términos que difieren apenas —e inesencialmente— de la definición gödeliana de función recursiva general presentada en la Sección 2.11.1. En este contexto, introduce la noción de conjunto recursivamente enumerable, que ha llegado a ocupar una posición central en la teoría de las funciones recursivas (o computables). Hoy día decimos que un conjunto $C \subseteq \mathbb{N}$ es *recursivamente enumerable* si y sólo si *C* es el dominio de una función recursiva parcial (en el sentido explicado al final de la p.

[10] Recuérdese que $\{F\}(m_1, \ldots, m_q)$ es una abreviatura. Por ejemplo, $\{F\}(m_1,m_2,m_3)$ es propiamente la fbf $\{\{\{F\}(m_1)\}(m_2)\}(m_3)$.

[11] La restricción a "enteros positivos" se explicó en la nota 8.

368). Un conjunto cualquiera es recursivamente enumerable si sus miembros tienen números de identidad y el conjunto de estos números es recursivamente enumerable. Estas definiciones son claras y simples, pero no dejan ver inmediatamente por qué dichos conjuntos se llaman de ese modo. En el escrito de Church, el concepto aparece bajo otra perspectiva. Una secuencia infinita de objetos cualesquiera, $a_0, a_1, a_2, \ldots$ puede verse como una aplicación con dominio $\mathbb{N}$ que asigna a cada número $n$ el $n$-ésimo término de la secuencia. Supongamos ahora que cada término $a_n$ tiene un número de identidad $\#a_n$. Diremos, con Church, que la secuencia $a_0, a_1, a_2, \ldots$ es recursiva si la función $n \mapsto \#a_n$ es recursiva. El conjunto $\{x|(\exists n \in \mathbb{N})(x = a_n)\}$ es *recursivamente enumerable* si la secuencia $a_0, a_1, a_2, \ldots$ es recursiva.[12] Obviamente, la función recursiva $n \mapsto \#a_n$ enumera el conjunto descrito. Para pasar de esta definición a la que formulé primero basta considerar lo siguiente: (i) si $f$ es una función recursiva total, sus valores, $f(0), f(1), f(2), \ldots$ forman una secuencia recursiva, de modo que el alcance de $f$, esto es, el conjunto $\{x|(\exists n \in \mathbb{N})(x = f(n))\}$ es un conjunto recursivamente enumerable, según nuestra segunda definición; (ii) si $f$ es una función recursiva total y $g$ es la función definida en el alcance de $f$ por la condición $g(x) = \mu y(f(y) = x)$, e indefinida fuera del alcance de $f$, se puede probar que $g$ es una función recursiva parcial; (iii) si $g$ es una función recursiva parcial, se puede probar que existe una función recursiva total $f$ cuyo alcance es el dominio de $g$. En virtud de (ii) y (iii) es claro que un conjunto numérico es el dominio de una función recursiva parcial si y sólo si es el alcance de una función recursiva total. Por lo tanto, las dos definiciones de conjunto recursivamente enumerable son equivalentes.

Church enuncia sin demostración una lista de doce teoremas que culmina con estos dos:[13]

12 La definición original de Church se aplica sólo a conjuntos de fbfs: "Un conjunto de fórmulas bien formadas se dice recursivamente enumerable si hay una secuencia recursiva infinita que consta exclusivamente de fórmulas del conjunto y contiene cada fórmula del conjunto por lo menos una vez" (1936, p. 96). Pero la generalización que di arriba es enteramente natural.

13 Church indica que el Teorema XVI fue demostrado por Kleene, e independientemente, en una forma ligeramente distinta, por Rosser, y que el Teorema XVII fue demostrado independientemente por Kleene y el propio Church, más o menos al mismo tiempo.

TEOREMA XVI. *Toda función recursiva es λ-definible.*
TEOREMA XVII. *Toda función λ-definible es recursiva.*

Llegado a este punto, Church está listo para dar la definición de calculabilidad efectiva anunciada al comienzo.

> Ahora definimos la noción, ya comentada, de una función *efectivamente calculable* de enteros positivos identificándola con la noción de una función recursiva de enteros positivos (o de una función λ-definible de enteros positivos). Pensamos que esta definición se justifica por las siguientes consideraciones, en la medida en que sea posible obtener una justificación positiva para la elección de una definición formal correspondiente a una noción intuitiva.
>
> (Church 1936, p. 356)

Junto con anunciar esta definición, Church anotó que ella está respaldada por el hecho de que los conceptos de función recursiva y de función λ-definible, de construcción tan diversa, tienen la misma extensión;[14] pero aquí aduce otras razones. Por una parte, recuerda que (como señaló en el pasaje que cité en la p. 371) cada función λ-definible tiene un algoritmo para calcular sus valores. Por otra parte, arguye que toda función calculable mediante un algoritmo tiene que ser λ-definible. Por ejemplo, si una función unaria $F$ es calculable mediante un algoritmo, tiene que haber *en alguna notación*, para cada $n \in \mathbb{N}$, una serie finita de expresiones $e_{n0}, e_{n1}, \ldots, e_{nr(n)}$, tal que:

(i) dado el número $n$, $e_{n1}$ sea efectivamente calculable,
(ii) dados $n$ y las expresiones $e_{nh}$ $(0 \leq h < k)$, $e_{nk}$ sea efectivamente calculable $(1 \leq k \leq r(n))$, y
(iii) cuando se ha construido la serie completa de expresiones hasta $e_{nr(n)}$, sea patente que el cálculo ha terminado y cuál es el valor de $F(n)$.

Si asignamos el gödel $\#[a]$ a cada expresión $a$ y acordamos que el gödel de la lista $\langle e_{n0}, e_{n1}, \ldots, e_{nk}\rangle$ sea el número $2^{\#[e_{n0}]} \times 3^{\#[e_{n1}]} \times \ldots \times p_k^{\#[e_{nk}]}$, pode-

14 "El hecho de que dos definiciones de calculabilidad efectiva tan distantes entre sí (*so widely different*) y, sin embargo, igualmente naturales (en opinión del autor) resulten equivalentes refuerza las razones aducidas más abajo para creer que ellas caracterizan dicha noción del modo más general que es compatible con nuestra habitual comprensión intuitiva de la misma" (Church 1936, p. 346, n. 3).

mos definir las funciones binarias $G$ y $H$ como sigue:

$G(n,x) = 0$, si $x$ es el gödel de $\langle e_{n0}, e_{n1}, \dots, e_{nr(n)}\rangle$;
$G(n,x) = \#[e_{nk}]$, si $k = h+1$, $h < r(n)$ y $x$ es el gödel de $\langle e_{n0}, e_{n1}, \dots, e_{nh}\rangle$;
$G(n,x) = 1$, en cualquier otro caso;
$H(n,x) = F(n)$, si $G(n,x) = 0$;
$H(n,x) = G(n,x)$, si $G(n,x) \neq 0$.

Si $G$ y $H$ son λ-definibles o recursivas, también lo es $F$. Como señala Church, no es fácil imaginarse en qué podría consistir la postulada calculabilidad de las series $e_{n0}, e_{n1}, \dots, e_{nr(n)}$, para cada $n \in \mathbb{N}$, si $G$ y $H$ no son λ-definibles o recursivas.

Sin embargo, Church no se da por contento con este argumento, y adelanta otro. Nos pide imaginarnos "un sistema particular de logica simbólica" que contenga un signo = para la igualdad entre números naturales, un signo { } ( ) para indicar la aplicación de una función unaria al respectivo argumento y numerales que representen unívocamente a los números naturales. Los teoremas del sistema comprenden una lista finita o numerable de expresiones, los *axiomas formales*, así como todas las expresiones derivables de éstas por una serie finita de aplicaciones de operaciones pertenecientes a una lista finita o numerable de *reglas procesales* (*rules of procedure*). "Para que el sistema sirva a todos los propósitos a que un sistema de lógica simbólica normalmente se destina es necesario que cada regla procesal sea una operación efectivamente calculable, que el conjunto completo de las reglas procesales (si es infinito) sea efectivamente enumerable y que la relación entre un entero positivo y la expresión que lo representa sea efectivamente determinable" (1936, p. 357). Hecha esta observación, Church nos invita a *suponer* que ella significa que, en términos de una asignación de gödels a las expresiones de la lógica, cada regla procesal es una operación *recursiva*, el conjunto de tales reglas es *recursivamente enumerable*,[15] el conjunto de los axiomas formales es *recursivamente enumerable* y la función que asigna a cada número el gödel del numeral respectivo es *recursiva*. Digamos con

15 "En el sentido de que existe una función recursiva $\Phi$ tal que $\Phi(n,x)$ es [el gödel] del resultado de aplicar la $n$-ésima regla procesal al conjunto ordenado finito de fórmulas [cuyo gödel] es $x$" (Church 1936, p. 357).

Church que una función unaria[16] $F$ es *calculable en la lógica* (*within the logic*) si hay en la lógica una expresión $f$ tal que $\{f\}(\mu) = \nu$ es un teorema si y sólo si $F(m) = n$ y $\mu$ y $\nu$ son los numerales que corresponden a $m$ y $n$, respectivamente. Como el conjunto completo de los teoremas es recursivamente enumerable, no cuesta nada probar que toda función unaria calculable en la lógica es recursiva (o $\lambda$-definible).

> Se ha mostrado así que no es posible obtener una definición de calculabilidad efectiva más general que la arriba propuesta por ninguna de las dos vías que se ofrecen naturalmente, a saber, (1) definir una función como efectivamente calculable si hay un algoritmo para calcular sus valores y (2) definir una función $F$ (de un entero positivo) como efectivamente calculable si, para cada entero positivo $m$ hay un entero positivo $n$ tal que $F(m) = n$ es un teorema demostrable.
>
> (Church 1936, p. 358)

Con estas palabras termina la presentación original de la llamada Tesis de Church. Me ha parecido oportuno repetir sus pormenores para que el lector se forme un juicio propio al respecto. Describirla como “definición” favorece un equívoco. En rigor, si la expresión ‘efectivamente calculable’ en adelante va a significar *por convención lingüística* lo mismo que ‘recursiva’ o ‘$\lambda$-definible’, no puede conservar su significado corriente, establecido mucho antes de que estos dos términos técnicos hicieran su aparición.[17] Gracias a que en castellano tenemos los dos sinónimos, ‘computable’ y ‘calculable’, he podido evitar ese equívoco, reservando ‘computable’ para los conceptos precisos —y equivalentes— introducidos en los años 30, y dejándole a ‘calculable’ su sentido habitual. Con esta terminología, los argumentos de Church pueden describirse así: se proponen dos análisis de lo que constituye un procedimiento de cálculo, basados, respectivamente, en el modo corriente de

[16] Como dice Church con toda razón, la ampliación de este concepto a funciones $n$-arias “es inmediata”.

[17] No cabe alegar que como la extensión de ese significado corriente es “vaga”, el matemático puede libremente deslindarla como le parezca oportuno. Si antes de 1928 alguien hubiera pretendido *definir* la expresión ‘efectivamente calculable’ *identificándola* con ‘pr-recursiva’, el descubrimiento de la función de Ackermann habría puesto en evidencia que esa definición no sirve. La ausencia de criterios precisos, enunciables de una vez por todas, para aplicar un concepto corriente (“intuitivo”) no significa que no sepamos reconocer cuando se lo usa bien o mal.

ejecución de algoritmos "con papel y lápiz" y en el "cálculo lógico" *à la* Frege-Russell; en ambos casos, la equivalencia *calculable* ⇔ *computable* se deduce del análisis, pero sólo después que cada uno de los pasos simples que, a la luz de ese análisis, constituyen el cálculo ha sido identificado con una operación de cómputo (recursiva o λ-definible). La equivalencia *global* de calculabilidad y computabilidad se deduce, pues, de su identidad postulada a nivel *local*. Por esto, el argumento no es propiamente circular, puesto que avanza de lo simple a lo compuesto, de los elementos al todo. Pero requiere que adoptemos como premisa, con respecto a los elementos de un procedimiento de cálculo, la identidad que se busca establecer como conclusión con respecto al procedimiento entero. En todo caso, gracias a los análisis de Church ha quedado muy claro que para convencernos de la falsedad de su Tesis habría que producir un algoritmo "de papel y lápiz" en que el tránsito de una línea a la próxima no pueda representarse mediante una función recursiva, o un cálculo lógico efectivo que admita una regla de inferencia no recursiva o tenga un conjunto de axiomas o de reglas de inferencia no recursivamente enumerable. Como en más de medio siglo nadie ha inventado un algoritmo o un cálculo lógico así, es razonable creer que no hay ejemplos contrarios a la Tesis de Church. Por otra parte, la misma índole del concepto corriente de calculabilidad no permite enunciar condiciones necesarias y suficientes para su aplicación; al darlas uno justamente *deja atrás* el concepto corriente —"natural" o, si se quiere, "histórico"— y lo reemplaza con uno artificial. Pero eso mismo impide ofrecer una demostración matemática de que el concepto corriente de función calculable coincide en extensión con el concepto artificial de función computable (recursiva, λ-definible). Por eso, no parece justo describir la Tesis de Church como una conjetura matemática que aguarda ser demostrada (comparable, digamos, a la conjetura de Legendre de que π es un número transcendental, antes de que fuera demostrada por Lindemann). La veo más bien como una *decisión* de aceptar la computabilidad como criterio de calculabilidad hasta que un contraejemplo fuerce a pensar de otro modo. Como todo el enorme incremento reciente de la capacidad humana de calcular (en el sentido ordinario) depende de la existencia de máquinas expresamente diseñadas para computar (en nuestro sentido técnico), no es verosímil que esa decisión vaya a revocarse.[18]

[18] Basándose en el análisis de la idea general de computación por Turing (*vide infra*,

La lista de teoremas que desemboca en la equivalencia entre funciones λ-definibles y recursivas también incluye los siguientes: VII. *El conjunto de las fbfs es recursivamente enumerable*. XI. *La relación de convertibilidad inmediata entre fbfs es recursiva*. XII. *Es posible asociar a cada fbf F una enumeración* $G_0, G_1, G_2, \ldots$ *de todas las formulas que cumplen la condición F conv* $G_k$ ($k \in \mathbb{N}$), *de tal modo que la función binaria* $\langle \#[F], k \rangle \mapsto \#[G_k]$ *sea una función recursiva*. XIII. *La propiedad de una fbf de estar en forma normal principal es recursiva*. XIV. *El conjunto de las fbfs que están en forma normal principal es recursivamente enumerable*. XV. *El conjunto de las fbfs que tienen una forma normal es recursivamente enumerable*. Algunos de estos teoremas se emplean en la demostración del siguiente resultado, que a su vez es la base de la demostración del Teorema de Church,

---

Sección 2.11.2) y la demostrada equivalencia de 'función recursiva' y 'función computable según Turing', Mendelson 1990 arguye que la Tesis de Church no es menos cierta que otras proposiciones básicas de la matemática que nadie sueña con llamar 'tesis'. Cita cuatro paralelos: (1) la definición habitual de 'función' (lo que en este libro llamamos 'aplicación') como una relación binaria $R$ tal que $\forall x \forall y \forall z(Rxy \wedge Rxz \rightarrow y = z)$; (2) la definición de 'verdad' propuesta por Tarski (*vide infra*, Capítulo 3.1); (3) la definición de 'validez lógica' en la teoría de modelos (también tratada en nuestro Capítulo 3.1), y (4) la definición de 'límite' por Weierstraß. Ninguno de estos paralelos me parece convincente. (1) El citado concepto de 'función' *amplía* resueltamente el concepto matemático que era corriente siglo XVIII, pero no cabe duda de que lo *incluye*; en cambio, quienes llaman 'tesis' a la definición de Church dirían que no sabemos con seguridad que todas las funciones que son calculables en la acepción corriente son λ-definibles o recursivas. (2) La definición de 'verdad' de Tarski se refiere a una propiedad de las fórmulas de un sistema formal y sólo una persona ignorante o muy irreflexiva puede entenderla como un intento de definir "rigurosamente" el concepto ordinario de verdad. (3) La definición modelista de 'validez lógica' se halla en el mismo caso, pero, claro está, no hay un concepto corriente con el que éste pudiera confundirse (o identificarse); por otra parte, contra lo que sugiere Mendelson, no goza de aceptación incuestionada entre los cultivadores de la lógica (cf. Etchemendy 1990). (4) Me parece que la existencia de alternativas al análisis clásico (Weyl 1918; Brouwer 1918, 1919, 1923; Lorenzen 1965; Bishop 1967; cf. asimismo Robinson 1961, 1966) y las sucesivas generalizaciones y redefiniciones del concepto de convergencia dentro de la misma tradición clásica (Moore y Smith 1922, H. Cartan 1937, Weil 1937) indican que el concepto de límite de Weierstraß no tiene el carácter definitivo que le atribuye Mendelson, aunque todavía se enseñe a los principiantes (con toda razón, como se les enseña la noción newtoniana de fuerza, igual a la masa invariable multiplicada por la aceleración absoluta).

que da una solución (negativa) al *Entscheidungsproblem* del cálculo predicativo de primer orden:

TEOREMA XIX. *No existe una función recursiva binaria f tal que (i) f(x,y)* = 0 *si x* = #[X], *y* = #[Y] *y* X *e* Y *son fbfs tales que* X *conv* Y, *y (ii) f(x,y)* = 1 *en cualquier otro caso*.

Church demuestra que el problema de hallar una función recursiva binaria *f* que cumpla las condiciones (i) y (ii) del Teorema XIX equivale al problema de hallar una función recursiva unaria *g* tal que *g*(*z*) = 0 *si z* = #[*Z*] y *Z* es una fbf que tiene una forma normal, y *g*(*z*) = 1 en cualquier otro caso. En vista de esto, el Teorema XIX es un simple corolario del siguiente

TEOREMA XVIII. *No existe una función recursiva unaria g tal que (i) g(x)* = 0 *si x* = #[X] *y* X *es una fbf que tiene una forma normal*, *y (ii) f(x)* = 1 *en cualquier otro caso*.

De este modo, el Teorema XVIII viene a ser la pieza clave en la solución del *Entscheidungsproblem*.[19] No voy a dar aquí una demostración detallada de este teorema, pero parafrasearé la explicación sumaria que Church mismo da para aclarar la suya. Supongamos que el Teorema XVIII es falso, y que existe la función recursiva *g* descrita en su enunciado. Sea *R* una fbf cualquiera. Si *g*(*R*) = 0, *R* tiene formas normales, una de las cuales es su forma normal principal *R**. Según el Teorema XII, hay una enumeración de todas las fbfs en las que *R* es convertible. Recorriéndola se encuentra, tarde o temprano, la forma normal principal *R**. *R* es convertible en un numeral si y sólo si *R** es un numeral (por cuanto (*a*) la forma normal principal de *R* es única y (*b*) cada numeral está en forma normal principal). Por lo tanto, si el Teorema XVIII es falso, es posible determinar mediante un procedimiento de cómputo si una fbf cualquiera es o no convertible en un numeral. Según el Teorema XV, hay una enumeración efectiva $A_0$, $A_1$, $A_2$, . . . de todas las fbfs que tienen una forma normal. Sea *n* el numeral correspondiente al número *n*. Definimos la función unaria *E* por la doble condición: (i) *E*(*n*) = 0

19 Con certero sentido de la jerarquía, Kleene (1967, pp. 242, 246, 282) usa la expresión 'Church's Theorem' —'el Teorema de Church'— para referirse al Teorema XVIII. Pero esa no es la nomenclatura habitual.

si $\{A_n\}(n)$ no es convertible en un numeral y (ii) $E(n) = m+1$ si $\{A_n\}(n)$ *conv* **m** y *m* es el numeral correspondiente a *m*. La función *E* es computable y, por ende, λ-definible mediante una fórmula que Church llama $\mathfrak{e}$. La demostración rigurosa del Teorema XVIII culmina, precisamente, con la construcción de esta formula $\mathfrak{e}$ y la prueba de que, si **n** es un numeral, $\{\mathfrak{e}\}(\mathbf{n})$ es convertible en un numeral (Church 1936, p. 361). Siendo así parte de fbfs que tienen una forma normal, $\mathfrak{e}$ tiene también ella misma una forma normal. Ahora bien, todas las fbfs que tienen una forma normal figuran en la lista $A_0$, $A_1$, $A_2$, . . . , pero $\mathfrak{e}$ no puede ser una de las fórmulas de esta lista, puesto que, según la definición de *E,* es imposible, para cada $n \in \mathbb{N}$, que $\{A_n\}(\mathbf{n})$ sea convertible en el numeral en que $\{\mathfrak{e}\}(\mathbf{n})$ es convertible.[20] La suposición de que el Teorema XVIII es falso lleva así a una contradicción.

El Teorema XVIII exhibe un ejemplo de función numérica no recursiva: la función *g* definida en su enunciado. Como el conjunto de las fbfs que tienen una forma normal es recursivamente enumerable (por el Teorema XV), el conjunto de las fbfs que *no* tienen una forma tal no puede serlo.[21] El conjunto de los gödel de este conjunto es, pues, un ejemplo de conjunto

20 Si $\{\mathfrak{e}\}(\mathbf{n})$ es convertible en el numeral correspondiente a 0, $\{A_n\}(\mathbf{n})$ no es convertible en un numeral; si $\{A_n\}(\mathbf{n})$ es convertible en un numeral **m**, $\{\mathfrak{e}\}(\mathbf{n})$ es convertible en el numeral siguiente a **m** y, por lo tanto, no es convertible en **m** (si una fbf es convertible en un numeral, ese numeral es único, puesto que cada fbf tiene a lo sumo *una* forma normal principal y todo numeral *es* una forma normal principal). Como puede advertirse, el paso central de la prueba del Teorema XVIII reproduce el patrón clásico del argumento diagonal. Específicamente, la función *E* se ha definido deliberadamente de tal modo que *ninguna* fbf $A_n$ —perteneciente a la lista $A_0$, $A_1$, $A_2$, . . . — antepuesta al numeral **n** —correspondiente a su posición en la lista— sea convertible en el mismo numeral en que es convertible la fbf $\{\mathfrak{e}\}(\mathbf{n})$, donde $\mathfrak{e}$ es la fbf que representa a *E* en el cálculo λ. Esta discrepancia garantiza que $\mathfrak{e} \neq A_n$, para cada $n \in \mathbb{N}$.

21 Sea ϕ la función recursiva que enumera las fbfs que tienen una forma normal. Si hubiera una función recursiva ψ que enumerase las fbfs que no tienen una forma normal, entonces mediante el examen de la secuencia ϕ(0), ψ(0), ϕ(1), ψ(1), ϕ(2), ψ(2), . . . podríamos decidir en un número finito de pasos si una dada fbf *A* tiene o no una forma normal. En efecto, *A* sería igual, digamos, al *r*-ésimo término de esa secuencia. Si $r = 2n$, $A = \phi(n)$ y por lo tanto tiene una forma normal; si $r = 2n+1$, $A = \psi(n)$ y por lo tanto no tiene una forma normal. De esto se infiere, por la Tesis de Church, que la función característica de la clase de fbfs que tienen una forma normal es recursiva. De hecho, construir una definición rigurosa de tal función recursiva, dadas ϕ y ψ, demandaría sólo un pequeño esfuerzo adicional. Pero su existencia contradiría al Teorema XVIII. Por lo tanto, no puede existir la función recursiva ψ.

numérico no recursivamente enumerable. El Teorema XVIII implica además, como ya dije, el Teorema XIX. De éste Church infiere sin dificultad que "el Entscheidungsproblem es insoluble en el caso de cualquier sistema de lógica simbólica que sea ω-consistente [ . . . ] y lo bastante poderoso como para admitir ciertos métodos de definición y prueba comparativamente simples" (1936, p. 363). Señala que en un sistema de esta clase se podrá expresar el aserto de que los números *a* y *b* son los gödel de fórmulas *A* y *B* tales que *A* es inmediatamente convertible en *B*. Como una conversión es una secuencia finita de fórmulas cada una de las cuales es inmediatamente convertible en la siguiente (si la hay), también cabe expresar en tal sistema —mediante una fórmula que llamaremos $\Psi(a,b)$— el aserto de que *a* y *b* son los gödel de fórmulas *A* y *B* tales que *A conv B*. Además, si *A conv B* y $a = \#[A]$ y $b = \#[B]$, la fórmula $\Psi(a,b)$ será deducible en el sistema mediante una prueba consistente en una cierta secuencia finita de conversiones inmediatas que lleva del numeral correspondiente a *a* al numeral correspondiente a *b*. Y si *A* no es convertible en *B*, la ω-consistencia del sistema implica que $\Psi(a,b)$ no es deducible. Ahora bien, si el *Entscheidungsproblem* del sistema admitiera una solución positiva, habría un algoritmo para determinar si $\Psi(a,b)$ es o no deducible. Ese algoritmo permitiría entonces determinar, para cualquier par de fórmulas *A* y *B*, si *A* es o no convertible en *B*. En otras palabras, dicho algoritmo permitiría calcular la función *f* descrita en el enunciado del Teorema XIX. Pero *f* no es recursiva y, por lo tanto, según la Tesis de Church, no puede calcularse mediante un algoritmo.

Aunque el razonamiento anterior emplea una terminología apropiada para hablar del cálculo λ de Church, puede fácilmente adaptarse a cualquier cálculo lógico de cierta clase muy amplia. Al final de su artículo, Church observa que "en particular, si el sistema de *Principia Mathematica* es ω-consistente, su Entscheidungsproblem es insoluble" (1936, p. 363); pero no entra en detalles. En su "Nota sobre el Entscheidungsproblem", publicada el mismo año, Church aplica el Teorema XIX específicamente al cálculo predicativo de primer orden de Hilbert y Ackermann (1928), concluyendo que "el caso general del Entscheidungsproblem [de dicho cálculo] es insoluble" (1936a, en Davis 1965, p. 114). En el Capítulo 2.7 consideramos dos formas del *Entscheidungsproblem*, según que el procedimiento algorítmico en cuestión concierna la *deducibilidad* o la *validez* de las fórmulas de un cálculo lógico. Como se recordará, distinguíamos entre *d*-decisión y *v*-decisión. En

el caso del referido cálculo de Hilbert y Ackermann, los dos problemas coinciden puesto que —en virtud del Teorema de Completud de Gödel (Capítulo 2.8)— una fórmula de ese cálculo es válida si y sólo si es deducible. A este respecto, Church hace una observación importante. Él ha demostrado constructivamente que no hay un algoritmo de *d*-decisión para dicho cálculo: no puede existir una función recursiva que asigne el valor 0 al gödel de cada fórmula deducible y el valor 1 al gödel de cada fórmula no deducible. Por lo tanto, en virtud del Teorema de Completud, tampoco hay un algoritmo de *v*-decisión. Pero la demostración del Teorema de Completud no es constructiva, de modo que la insolubilidad del problema de la *v*-decisión no se ha establecido constructivamente. Por esta razón, Church concluye que "la insolubilidad de esta segunda forma del Entscheidungsproblem del [cálculo predicativo de primer orden de Hilbert y Ackermann] no puede considerarse incuestionablemente establecida" (1936a, en Davis 1965, p. 115).

### *2.11.3 Las máquinas de Turing*

El artículo de Alan Turing, "Sobre números computables, con una aplicación al Entscheidungsproblem" (1936), aborda el tema del cálculo efectivo desde un nuevo punto de vista, que lleva a una definición inesperadamente simple de función computable, equivalente a las que hemos considerado hasta aquí. Una definición esencialmente idéntica a la de Turing fue propuesta independientemente por Emil Post (1936).[22] Por su sencillez, esta definición de computabilidad es la preferida por los autores de obras didácticas. Aquí la utilizaré para demostrar, finalmente, el Teorema de Church (en la Sección 2.11.5). Para mayor claridad, basaré el razonamiento en una *variante* de la definición original de Turing. Pero antes voy a presentarla en los términos

[22] El artículo —muy conciso— de Post apareció un poco antes que el de Turing, con una nota que reconoce la prioridad temporal de éste y aclara que los autores trabajaron con total independencia. Ambos, por cierto, conocían la labor de Gödel y Church. En el trabajo de Post falta completamente la idea de una *máquina* computadora. Lo que Turing (1936) llama la configuración interna de la máquina —"el programa", como decimos hoy— Post (1936) lo presenta como instrucciones que ha de seguir un calculista (humano).

empleados por él en 1936.

Turing, como Church, usa el mismo vocablo para referirse a la noción corriente que le interesa deslindar con precisión y a la noción precisa definida por él. Pero en vez de 'calculable' —como decía Church— Turing dice 'computable'.[23] Ocasionalmente encierra esta palabra entre comillas, para indicar que se refiere a la noción precisa. Por mi parte, seguiré usando 'calculable' para la noción corriente y 'computable' para la noción técnica. Diré 'T-computable' cuando sea necesario distinguir el concepto de Turing (o sus variantes) de otras nociones equivalentes, como 'recursiva general' o 'λ-definible'.

El artículo de Turing se refiere expresamente a los *números computables*, esto es, "los números reales cuya expresión digital es calculable con medios finitos".[24] Pero, como Turing se apresura a señalar, su tratamiento se puede adaptar sin dificultad a "las funciones computables de una variable entera, o de una variable real o computable, los predicados computables, etc." (1936, p. 230). Con la variante arriba aludida aplicaré la noción de 'T-computable' a las funciones numéricas (aplicaciones de $\mathfrak{D} \subseteq \mathbb{N}^r$ en $\mathbb{N}$) de que tratan las dos subsecciones precedentes (las cuales, por lo demás —como señalé al comienzo de el Capítulo 2.11— entran en el cálculo de cualquier "número computable"). Pero primero atendamos a la noción original.

Números computables, en el sentido descrito, son precisamente los que calculaban los calculistas (*computers*, en inglés), esos abnegados trabajadores que, antes del advenimiento de la computadora electrónica, aplicaban "a mano" los algoritmos apropiados para construir tablas de logaritmos, senos, cosenos, etc., evaluar integrales, sumar series, al servicio de la ciencia y la

23 Reserva 'calculable' —más precisamente, 'effectively calculable'— para referirse al concepto de Church. Turing 1936 muestra en un apéndice que el concepto de Church tiene la misma extensión que el concepto de computabilidad definido por él.

24 Donde pongo "expresión digital", Turing escribe "decimal expression", pero lo que quiere decir con eso es 'expresión mediante dígitos de un sistema numérico cualquiera (fijo)'. Como veremos, sus "máquinas" —prefigurando la computadora electrónica— emplean el sistema binario. A propósito de ellas, Turing habla de "binary decimals". Ello tal vez no molesta en una lengua en que el nombre del 10 no proviene del latín *decem*; pero a nosotros un 'decimal binario' nos suena disparatado y preferimos hablar de 'dígitos binarios'.

ingeniería. Turing busca una definición precisa del término 'calculable con medios finitos'. Para eso, describe una clase de máquinas ideales concebidas —según él— de modo que simulen perfectamente las operaciones de un calculista. Llamaré *máquina de Turing* a cualquier objeto de esta clase. Un número real es "computable" si hay una máquina de Turing que lo computa. Después de explicar como operan sus máquinas, Turing introduce la idea que inspirará la invención de la computadora moderna: la *máquina de Turing universal* $\mathcal{U}$, que computa el número computado por cualquier máquina de Turing $\mathcal{M}$ cuando el número de identidad de $\mathcal{M}$ se registra como dato inicial en la memoria de $\mathcal{U}$. Siguiendo a Turing, daré primero la descripción general de una MT, luego ciertos resultados importantes de su trabajo, y por último repetiré el argumento en que basa su afirmación de que todo lo que pueda calcular un calculista es computable por una máquina de Turing.

Una máquina de Turing $\mathcal{M}$ se caracteriza, ante todo, por una lista finita $q_0, \dots, q_m$ de estados (Turing dice "$\mathcal{M}$-configurations") en que $\mathcal{M}$ puede encontrarse y un alfabeto finito $S_0, \dots, S_n$ de signos ("symbols") que $\mathcal{M}$ puede reconocer e imprimir. Uno de éstos, digamos $S_0$, es sólo un espacio en blanco. $\mathcal{M}$ lee y escribe en una cinta de papel cuadriculado . Entendemos que la cinta tiene una orientación; evocando nuestros propios hábitos de lectura y escritura, diré que $\mathcal{M}$ 'avanza' hacia la 'derecha' y 'retrocede' hacia la 'izquierda'. En un dado momento, cada cuadrado de la cinta contiene un solo signo (posiblemente, el blanco $S_0$). Suponemos que la cinta es infinita o, al menos, que le "crece" un cuadrado adicional en blanco apenas $\mathcal{M}$ llega a leer uno de sus extremos (éste es el único rasgo ideal de las máquinas de Turing, inimitable por calculistas y computadoras). Suponemos además que en el momento inicial hay a lo sumo un número finito de cuadrados que "no están en blanco", esto es, que contienen un signo diferente de $S_0$, (este supuesto mitiga lo irrealista del anterior). En cada momento, $\mathcal{M}$ está en uno de sus estados, leyendo un cuadrado de la cinta. Ese estado y el signo que hay en ese cuadrado determinan la próxima acción de $\mathcal{M}$, la cual consiste en ejecutar una de cada una de las tres alternativas siguientes: (i) escribir uno de los signos $S_0, \dots, S_n$ en sustitución del signo que lee, (ii) avanzar un cuadrado (llamemos a esto, $A$), retroceder un cuadrado ($R$) o permanecer detenida leyendo el mismo cuadrado ($D$) y (iii) pasar a uno de

los estados $q_0, \ldots, q_m$.[25] El funcionamiento de $\mathcal{M}$ queda, pues, completamente caracterizado por una lista finita de quíntuplos $\langle q_h,S_v,S_w,C,q_k\rangle$, tales que $1 \leq h,k \leq m$, $1 \leq v,w \leq n$, $C$ es uno de los tres caracteres $A$, $R$, $D$, y no hay dos quíntuplos que comiencen con el mismo par $\langle q_h,S_v\rangle$.[26] El quíntuplo $\langle q_h,S_v,S_w,C,q_k\rangle$ indica lo que $\mathcal{M}$ hace si lee el signo $S_v$ cuando se encuentra en el estado $q_h$: (i) reemplaza $S_v$ por $S_w$, (ii) efectúa la operación representada por $C$ y (iii) pasa al estado $q_k$. Diré que la lista de quíntuplos característica de $\mathcal{M}$ es el *programa* de $\mathcal{M}$ (Turing lo llama "la configuración completa" de $\mathcal{M}$, pero me parece que el término que he elegido no sólo es más breve, sino también, hoy por hoy, más expresivo). Para mayor precisión, supongo que el programa de una máquina de Turing está siempre ordenado alfabéticamente ($\langle q_h,S_v,S_w,C,q_k\rangle$ precede a $\langle q_f,S_u,S_w,C,q_k\rangle$ si $h < f$, o si $h = f$ y $v < u$).[27]

Turing propone un método sencillísimo para darle un número de identidad —un gödel— a cada máquina de Turing $\mathcal{M}$. Si escribimos los subíndices numéricos de sus estados y signos como listas de palotes (|), separamos con punto y coma (;) los quíntuplos que forman el programa de $\mathcal{M}$ y eliminamos

25 Obsérvese que la alternativa (i) incluye la posibilidad de dejar intacto el signo leído, puesto que es uno de los signos permitidos como sustituto. La alternativa (iii) incluye la posibilidad de seguir en el mismo estado.

26 En otras palabras, el programa de $\mathcal{M}$ es el grafo de una aplicación cuyo dominio está incluido en $\{0, \ldots, m\} \times \{0, \ldots, n\}$ y cuyo alcance está incluido en

$$\{0, \ldots, n\} \times \{A, R, Q\} \times \{0, \ldots, m\}.$$

27 La máquina $\mathcal{M}$ descrita es lo que Turing (1936, p. 232) llama una *máquina automática* o *a*-máquina, porque su funcionamiento está completamente determinado por el programa. Menciona además lo que llama una "máquina a elección" (*choice machine*) y que hoy llamamos una *máquina de Turing con oráculo* (MTO), cuyo funcionamiento no está completamente determinado por el programa. Éste indica que cuando la máquina está en ciertos estados leyendo ciertos signos hay que consultar una fuente externa u "oráculo"—un operador humano, una tabla de números aleatorios, una ruleta, etc.— para determinar lo que la máquina ha de hacer en esa situación. Así, si $\mathcal{M}_0$ es una MTO, el programa de $\mathcal{M}_0$ consta en parte de quíntuplos como los arriba descritos, en parte de triples de la forma $\langle q_h,S_v,?\rangle$ donde el signo de interrogación indica que hay que consultar el oráculo. La respuesta de éste tiene que ser un triple de la forma $\langle S_w,C,q_k\rangle$, pero no está predeterminado cuál será ese triple en cada caso. (Evidentemente, el programa de $\mathcal{M}_0$ puede fijar distintas listas finitas de respuestas posibles en las diversas situaciones en que prevé una consulta al oráculo; también puede contemplar varios oráculos, de modo que la respuesta de uno remita a veces a otro).

los paréntesis angulares y las comas, podemos representar ese programa inequívocamente mediante un número construido así: ponemos 1 en vez de un palote, 2 en vez de $S$, 3 en vez de $q$, 4 en vez de $A$, 5 en vez de $R$, 6 en vez de $D$ y 7 en vez de punto y coma. Por ejemplo, si $\mathcal{M}_1$ es una máquina de Turing con cinco estados y un alfabeto de dos signos cuyo programa, escrito en la forma acordada, es

$$\ulcorner q\ S_{|}\ S\ D\ q_{|};\ q_{|}\ S\ S\ A\ q_{||};\ q_{||}\ S\ S_{|}\ D\ q_{||||};\ q_{||}\ S_{|}\ S\ D\ q_{|||};\ q_{|||}\ S\ S\ A\ q_{|}\urcorner,$$

su gödel $\#[\mathcal{M}_1]$ es el número[28]

321.263.173.122.431.173.112.216.311.117.311.212.631.117.311.122.431.

La descripción anterior es bastante general y se aplica lo mismo a las máquinas de Turing que evalúan números computables, en el sentido de Turing, que a las que computan el número natural que una determinada función recursiva asigna a cada $r$-tuplo de números naturales, etc. Basaré en ella la definición variante anunciada arriba. Pero antes de proponerla, completemos la descripción original de una máquina de Turing capaz de computar la expresión digital de un número real o MTR (como diré para abreviar). Si se adopta, con Turing, el sistema numérico binario, una MTR requiere un alfabeto de sólo dos dígitos, 0 y 1, además del blanco $S_0$. Pero la programación se facilita si hay también signos auxiliares, que puedan usarse para "llevar cuentas" durante la computación (por ejemplo, para marcar el punto de la cinta en que la máquina debe interrumpir su avance y empezar a retroceder). Turing (1936, p. 118) asigna a sus MTRs un alfabeto con dos clases de signos, aparte del blanco: los dígitos 0 y 1, y signos auxiliares "de segunda clase". Sea, pues, $\mathcal{M}$ una MTR con un alfabeto así. He aquí como Turing concibe su funcionamiento: $\mathcal{M}$ empezará a operar en cierto estado $q_0$ leyen-

28 El lector puede entretenerse en comprobar que $\mathcal{M}_1$ hace lo siguiente cuando trabaja con una cinta vacía u ocupada sólo por una fila consecutiva y finita de palotes. Si $\mathcal{M}_1$ empieza a funcionar en el estado $q_0$ leyendo el primer palote de la fila, o un blanco cualquiera si la cinta está vacía, entonces, la cinta cesa de funcionar (i) leyendo un blanco en una cinta vacía si el número inicial de palotes era par y (ii) leyendo el único palote de una cinta por lo demás vacía si el número inicial de palotes era impar. Si entendemos que una fila de $n$ palotes consecutivos ($n \geq 0$) representa al número natural $n$, es claro que $\mathcal{M}_1$ computa la función característica de la clase de los impares (si esta función se define como en la p. 43; si la función característica se define —como en la p. 333— a la manera de Gödel, $\mathcal{M}_1$ computa la de los pares).

do un cuadrado cualquiera de una cinta enteramente en blanco; siguiendo las instrucciones de su programa, imprimirá dígitos y signos auxiliares en la cinta. Aunque Turing no lo dice, vamos a suponer que $\mathcal{M}$ nunca imprime un dígito en un cuadrado a la izquierda del cuadrado inicial. Bajo este supuesto, es claro que, si se ignoran los cuadrados en blanco u ocupados por signos que no son dígitos, los dígitos impresos por $\mathcal{M}$ forman una secuencia (finita o indefinidamente creciente). Turing la llama "la secuencia computada por la máquina". Si la secuencia computada por $\mathcal{M}$ es finita, diremos con Turing que $\mathcal{M}$ es una MTR *circular*. Obviamente, $\mathcal{M}$ es circular si y sólo si hay un momento a partir del cual se estaciona para siempre en un cuadrado o se sigue moviendo pero deja de escribir dígitos. Si $\mathcal{M}$ no es circular, la secuencia (infinita) computada por ella es la mantisa —expresada en el sistema binario— de un número real $\alpha$ en el intervalo [0,1].[29] Turing llama a $\alpha$ "el número computado por $\mathcal{M}$". Turing proclama que "un número es computable si la diferencia entre él y el número computado por una máquina no circular es un entero" (1936, p. 233).[30] Si 'computable' significa aquí *calculable* (en el sentido corriente), esta declaración enuncia una "Tesis de Turing" homóloga a la Tesis de Church.

Este concepto de circularidad entra en la demostración original de que hay una máquina de Turing *universal*, que Turing caracteriza como una MTR que, aplicada a una cinta que contiene el gödel de cualquier MTR no circular, computa la secuencia computada por ésta. También determina la formulación —y demostración— del resultado más importante de Turing 1936. Tal como allí se presenta, éste se refiere al "problema de la circularidad", que puede enunciarse así: determinar mediante un algoritmo si una MTR cuyo programa se conoce es o no circular. Turing demuestra que no puede existir una máquina de Turing que compute la función característica de la clase de las MTRs no circulares. Combinado con la tesis de que todo lo que es calculable es T-computable, este resultado da una solución negativa al problema

29 Recuérdese que la mantisa de un número real es la parte de su expresión digital que se escribe a la derecha de la coma (del punto en el sistema anglosajón utilizado por Turing).

30 Si hay una MTR que computa cierto número real $\alpha \in [0,1]$, no cuesta nada programar una MTR circular que escriba la parte entera de un número real $\beta$ tal que $|\alpha - \beta| \in \mathbb{N}$, y una tercera MTR (cuyo alfabeto incluya, posiblemente, la coma y el signo menos) que combine el producto de las dos anteriores.

de la circularidad. Como las MTRs nos interesan aquí muchísimo menos que las máquinas de Turing que computan funciones numéricas, es preferible estudiar la idea de la máquina de Turing universal en relación con éstas, así como el problema homólogo al de la circularidad que se plantea con respecto a ellas: determinar mediante un algoritmo si una máquina de Turing cuyo programa se conoce computa o no una función numérica. Este problema se llama habitualmente "el problema de la detención" (*the halting problem*) porque se lo puede formular asi: si $x$ es un número ¿es $x$ el gödel de una máquina de Turing que, aplicada a cualquier dato numérico de cierto tipo (que represente un $r$-tuplo para algún $r$ fijo), acaba deteniéndose ante un resultado numérico en la posición prescrita para recibirlo? En caso afirmativo, la máquina de Turing †[$x$] computa la función que asigna ese resultado al dato en cuestión. En caso negativo, esto es, si †[$x$] no se detiene nunca o se detiene en una posición distinta de la prescrita para recibir un resultado, obviamente †[$x$] no computa una función numérica. (Doy una explicación más exacta del "problema de la detención" en la Sección 2.11.5).

A continuación doy una descripción canónica de las máquinas de Turing que computan funciones numéricas. *Usaré la abreviatura* MT *para designar a las máquinas de Turing que obedecen a esta descripción*. En primer lugar, estipularemos que todas las MT tienen un mismo alfabeto y siguen un procedimiento convencional uniforme para la recepción de datos y entrega de resultados. El alfabeto consta solamente de dos signos: $S_0$, el blanco, que llamaré $B$, y $S_1$, el palote |. (Para un trabajo práctico de programación sería *cómodo* disponer además de unos pocos signos auxiliares, pero su adopción no expande el reino de las funciones computables). Una fila de palotes, impresos en cuadrados consecutivos de la cinta, precedida y seguida de un blanco, representa un número natural determinado. Como, según la convención seguida generalmente en este libro, el primer número natural es 0, representamos el número $n$ con una fila de $n$ + 1 palotes. Representamos el $r$-tuplo $\langle n_1, \ldots, n_r \rangle$ mediante $r$ filas de $n_1$+1, . . . , $n_r$+1 palotes, respectivamente, con un solo blanco entre cada dos filas consecutivas. Para mayor brevedad, llamaré **número* a la representación de un número mediante una fila de palotes, **r-tuplo* a la representación de un $r$-tuplo de números mediante una fila de $r$ filas de palotes. Digo que una MT *se aplica* a un *número cuando lee su primer palote, y que *se aplica* a un *$r$-tuplo cuando se aplica a su primer *número. Digo que un blanco *cierra* un *número o un *$r$-

tuplo cuando sigue inmediatamente a su último palote y que lo *abre* cuando precede inmediatamente a su primer palote. Gracias a la sencillez del alfabeto, podemos simplificar los programas. Arriba postulamos que una máquina de Turing reacciona en un cierto estado ante un signo dado con *tres* operaciones: (i) reemplazar el signo leído con un signo dado (posiblemente el mismo), (ii) permanecer ante el mismo cuadrado o avanzar o retroceder a un cuadrado adyacente, (iii) adoptar cierto estado (posiblemente el mismo). Ahora supondremos que la reacción de una máquina capaz de $m$ estados diferentes consiste en ejecutar una de cada una de las *dos* alternativas siguientes: (i) reemplazar el signo leído con $B$ o con |, o avanzar un cuadrado ($A$), o retroceder un cuadrado ($R$);[31] (ii) adoptar el estado $q_k$ ($0 \leq k < m$). El programa está dado entonces por una lista de cuádruplos de la forma $\langle q_h,S_v,C,q_k\rangle$, donde $1 \leq h,k < m$, $v \in \{0,1\}$, $C$ es uno de los cuatro caracteres $B$, |, $A$, $R$, y no hay dos cuádruplos que comiencen con el mismo par $\langle q_h,S_v\rangle$.[32] Supondremos que una MT dada siempre empieza a operar en cierto estado, el *estado inicial* de esa MT, que siempre llamaré $q_0$. Obviamente, MT sólo llega a detenerse si su lista de estados incluye por lo menos un *estado final*, esto es, un estado que figura en cuarto lugar en uno o más cuádruplos del programa (y así hay operaciones que conducen a él), pero no figura en primer lugar en ningún cuádruplo del programa (y, por ende, no es el punto de partida de ninguna operación). Finalmente, supondremos que una MT nunca retrocede más de dos cuadrados a la izquierda de su posición inicial.

Sea $\mathcal{M}$ una MT. Diremos que $\mathcal{M}$ computa la función $r$-aria $f:\mathbb{N}^r \to \mathbb{N}$ si su programa determina el siguiente comportamiento: cuando $\mathcal{M}$ se aplica en

31 *Nota sobre la notación*. Hablando de máquinas de Turing, uso en adelante varias formas de las letras *a* y *r* (mayúscula, minúscula, cursiva, recta, helvética, griega, gótica) para referirme a objetos u operaciones relacionados, respectivamente, con las dos direcciones de la cinta: *avance* (hacia la derecha) y *retroceso* (hacia la izquierda). En inglés y alemán, usan la *r* (de *right*, *Rechte* = derecha) en lugar de nuestra *a*, la *l* (de *left*, *Linke* = izquierda) en vez de nuestra *r*. No quise usar el par de letras *d* e *i*, porque la *i* minúscula tradicionalmente denota un *índice* y la mayúscula se parece demasiado al 1 y al palote |.

32 Al asignar un gödel a uno de estos programas, hay que distinguir el palote que es un signo del alfabeto, del palote que utilizamos como subíndice numérico para identificar a cada estado. Escribiendo los cuádruplos como antes, separados por punto y comas, podemos asignar el dígito 1 al palote subíndice (como antes), el 2 al palote signo, el 0 al blanco, el 3 a *q*, el 4 a *A*, el 5 a *R* y el 6 al punto y coma.

su estado inicial al *$r$-tuplo $\langle n_1, \ldots, n_r \rangle$, escrito en una cinta por lo demás en blanco, $\mathcal{M}$ se detiene, al cabo de un número finito de operaciones, aplicada al *número $f(n_1, \ldots, n_r)$ en una cinta por lo demás en blanco. En otras palabras, si $\mathcal{M}$ recibe una representación de $\langle n_1, \ldots, n_r \rangle$ como dato, entrega una representación de $f(n_1, \ldots, n_r)$ como resultado. En lo sucesivo, suelo llamar $\mathcal{M}_f$ a la MT que computa cierta función $f$. Si $f$ es una función parcial (definida en una parte propia de $\mathbb{N}^r$), la máquina que la computa tiene el comportamiento descrito cuando el contenido inicial de la cinta representa un $r$-tuplo en que $f$ está definida, pero sigue operando sin detenerse jamás si dicho contenido representa un $r$-tuplo en que $f$ no está definida.

De acuerdo con estas convenciones, el programa $\ulcorner q_0B|q_1;\ q_0|Aq_0;\ q_1BAq_2;\ q_1|Rq_1 \urcorner$ define una MT que computa la función sucesor, $x \mapsto x+1$. El lector puede persuadirse de esto, ya sea ejecutando el programa, por ejemplo, en una cinta en que está escrito el *número 3, ya sea estudiando la representación gráfica del programa que presento y explico al comienzo de la Sección 2.11.4 (p. 393).

En esa sección analizo los programas de otras MT elementales y enseño a descomponerlos en partes que pueden combinarse para formar nuevos programas. Los recursos presentados allí me ayudarán a probar en la Sección 2.11.5 que toda función recursiva es T-computable. Probaré también que toda función T-computable es recursiva y utilizaré este resultado para demostrar, a propósito de las MT que computan funciones numéricas, (I) que hay una MT universal, esto es, una MT que computa la función $\langle \#[\mathcal{M}_f], x \rangle \mapsto f(x)$;[33] y (II) que el problema de la detención es insoluble, esto es, que no hay una MT que compute la función característica del conjunto de números $\{x{:}x$ es el gödel de una MT que computa una función numérica$\}$. Pero antes de abordar estos temas, veamos cómo razonaba Turing para convencerse —y convencernos— de que sus máquinas son capaces de simular perfectamente las operaciones de un calculista humano.

El argumento de Turing incluye (i) un “llamado a la intuición”, esto es, a nuestra idea de lo que puede o no puede hacer un calculista; (ii) una definición precisa de computabilidad —distinta de las presentadas aquí— que resulta ser equivalente a la T-computabilidad, y (iii) una colección de ejem-

33 Aplicada a una cinta que contiene el gödel de la MT que computa la función $f$ seguido del *número $x$, la MT universal produce el valor de $f$ correspondiente al argumento $x$.

plos de “vastas clases de números que son computables”. Como las definiciones de computabilidad conceptualmente diversas pero extensionalmente equivalentes se han multiplicado desde 1936, si (ii) ya poseía alguna fuerza persuasiva entonces, ahora ella tiene que ser arrolladora. En la Sección 2.11.5 probaré que una función numérica es T-computable si y sólo si es recursiva. Omitiré, por eso, la parte (ii) del argumento de Turing. También la parte (iii) se ha vuelto prescindible, puesto que en medio siglo nadie ha podido señalar una clase de números —o una función numérica— que sea calculable pero no computable (en particular, en virtud del diseño mismo de las computadoras electrónicas, todo número o función calculable por una de ellas tiene que ser T-computable). En cambio, la parte (i) merece nuestra atención por su gran originalidad y sencillez, porque pone de manifiesto las consideraciones que inspiraron el diseño de las máquinas de Turing y porque convenció a Gödel de que los nuevos conceptos precisos de computabilidad capturaban la noción ordinaria de algoritmo.[34]

Turing recuerda que los calculistas normalmente trabajan escribiendo signos en un papel. “Podemos suponer que ese papel está cuadriculado como el cuaderno de aritmética de un niño. En la aritmética elemental suele aprovecharse el carácter bidimensional del papel. Pero ello es prescindible y pienso que todos estarán de acuerdo en que la bidimensionalidad del papel no es un requisito esencial para calcular. Supongo, pues, que el cálculo se lleva a cabo en papel unidimensional, esto es, en una cinta dividida en cuadrados” (1936, p. 249).[35] Cada cuadrado y los signos que pueden escribirse en él son, por cierto, finitos. Por eso hay que suponer que el calculista sólo tiene un número finito de signos diferentes a su disposición. En efecto, un alfabeto infinito de signos legibles tendría que incluir figuras tan poco diferenciadas que ningún calculista sería capaz de distinguirlas. Turing se apresura a señalar que la restricción del número de signos no es grave, puesto que siempre se puede usar una fila de signos como si fuera un signo más (así, en nuestro CP1 generamos infinitas variables con los dos signos $x$ y |, etc.). La conducta de un calculista *está determinada* en cada momento por los signos a que presta

[34] Gödel, CW, I, 369; OC, p. 196; citado en la Sección 2.11.1.

[35] ¿Fijó Turing con esta decisión el carácter distintivo de la computadora electrónica, que ejecuta sus operaciones en serie y hasta el día de hoy (1 de abril de 1998) no era capaz de trabajar “en paralelo”, como lo hace el cerebro humano?

atención y por su “estado mental” (*“state of mind”*). Estos determinan además el estado mental del calculista en el próximo momento. Turing apunta aquí a un carácter esencial del concepto de algoritmo: si el calculista no ajusta cada una de sus operaciones —conforme a las reglas del cálculo— a la situación alcanzada anteriormente cometerá un error o interrumpirá su tarea; y la situación —de la cinta y de su mente— después de una operación depende exclusivamente de ésta y de los factores que la determinan.[36] Hay que suponer, además, que el número de signos que el calculista puede observar en un momento dado no excede un cierto número *B*. Para observar más de *B* signos, tiene que hacer observaciones sucesivas. Turing también supone finito el número de estados mentales de que el calculista es capaz, porque, según él, si fuesen infinitos habría algunos tan parecidos que el calculista no sabría distinguirlos. Aunque Turing obviamente no puede probar que la variedad de la conciencia de sí esté sujeta a esta limitación,[37] creo que anota aquí acertadamente un rasgo esencial de esa forma de actividad que llamamos ‘calcular’. Por otra parte, como Turing se apresura a señalar, esta limitación puede obviarse escribiendo más signos en la cinta.

> Imaginemos que las operaciones ejecutadas por el calculista se han dividido en “operaciones simples” que son tan elementales que no es fácil imaginar una división ulterior. Cada una de estas operaciones consiste en cierta alteración del sistema físico formado por el calculista y su cinta. Conocemos el estado del sistema si conocemos la secuencia de signos en

36 Hay que suponer, claro, que se trata de la situación *pertinente*, esto es, del estado de la cinta y de la mente del calculista *en cuanto es significativa para el procedimiento de cálculo*. Si, al tiempo que escribe el próximo símbolo, el calculista derrama una taza de café sobre el papel, la nueva situación del papel y de su mente no depende exclusivamente de la operación de cálculo ejecutada. Aunque parece idiota, esta observación es importante, pues indica que para deslindar en el acontecer real lo que es o no es pertinente a un procedimiento de cálculo hay que tener más o menos en claro qué entendemos por ‘procedimiento de cálculo’.

37 Aunque identifiquemos, como ahora está de moda, la mente con el encéfalo o una parte de él, siempre podría asumir infinitos estados diferentes en el trascurso de una vida, puesto que el espacio y el tiempo son infinitamente divisibles. En cuanto a la aptitud de la conciencia humana para distinguirlos, sabemos demasiado poco para aventurar una conclusión al respecto. Distinto es el caso de los símbolos que son marcas de tinta en un papel, puesto que podemos señalar umbrales bajo los cuales dos marcas son indiscernibles a simple vista o bajo un microscopio.

> la cinta, cuáles de ellos observa el calculista (posiblemente en un cierto orden) y el estado mental del calculista. Cabe suponer que en una operación simple no se altera más de un signo. Cualquier otra alteración puede dividirse en cambios simples de esta clase. [ . . . ] Cabe suponer, sin pérdida de generalidad, que los cuadrados cuyos signos se alteran son siempre cuadrados "observados".
>
> Además de la alteración de signos, las operaciones simples tienen que incluir cambios en la distribución de los cuadrados observados. Los nuevos cuadrados observados tienen que ser inmediatamente reconocibles para el calculista. Me parece razonable suponer que sólo pueden ser cuadrados cuya distancia del cuadrado más próximo de los observados recién no exceda cierta magnitud fija. Digamos pues que cada uno de los nuevos cuadrados observados dista $L$ cuadrados o menos de uno de los que acaban de observarse.
>
> (Turing 1936, p. 250)

Turing contempla también la posibilidad de que algunos cuadrados estén marcados, lo que permitiría al calculista encontrarlos en cualquier momento aunque disten más de $L$ cuadrados de cualquiera de los cuadrados bajo observación. Señala que, si cada uno está marcado con un signo, podemos agregarlos a la colección de cuadrados observada en cada momento sin que ésta deje de ser finita. Por otra parte, si están marcados con secuencias de signos de longitud arbitraria su reconocimiento no puede ocurrir en un sólo acto, sino que requiere operaciones sucesivas. Así, por ejemplo, para distinguir las secuencias

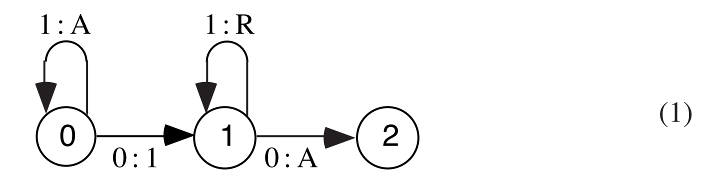
1111111111111111111111111111111 y 111111111111111111111111111111

hay que compararlas dígito por dígito, "tal vez marcando los dígitos con un lápiz para asegurarse de que uno no los cuenta dos veces".

Sin duda, una máquina de Turing puede ejecutar la labor descrita. "A cada estado mental del calculista corresponde un estado (*m-configuration*) de la máquina" (1936, p. 251). Cuando el calculista ve $B$ cuadrados a la vez, la máquina los lee consecutivamente. Llamemos a esto una *etapa* de su funcionamiento. En cada etapa $E$, la máquina puede alterar un signo en uno de los cuadrados leídos en $E$ o moverse de un cuadrado leído en $E$ a un cuadrado cualquiera que diste a lo sumo $L$ cuadrados de alguno de los cuadrados leídos en $E$. La operación ejecutada y el estado al que la máquina pasa dependen del signo leído y del estado en que la máquina se encontraba.

### *2.11.4 Diagramas y ejemplos*

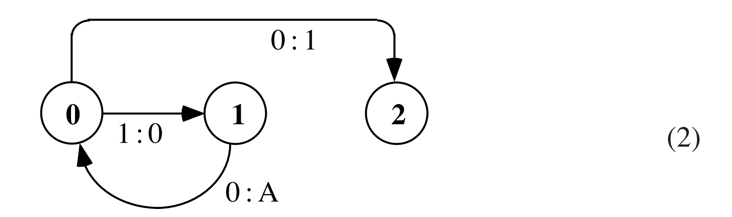
Di arriba el programa de una MT que computa la función sucesor $x \mapsto x+1$. El diagrama siguiente —leído como explicaré enseguida— da una idea más clara de su forma de operar.




Los tres círculos representan los tres estados posibles de la MT, identificados por su respectivo índice numérico (en negrita). Para mayor claridad, llamo 0 al signo $S_0$ (el blanco) y 1 al signo $S_1$ (el palote). A y R denotan las operaciones de avanzar y retroceder un cuadrado. Como puede observarse, cada flecha se origina en un estado, va acompañada de dos caracteres (separados por dos puntos), y apunta a un estado. Estos cuatro elementos constituyen obviamente un cuádruplo del programa. Cada flecha debe entonces entenderse así: cuando la MT está en el estado donde la flecha se origina, leyendo el signo anotado a la izquierda de los dos puntos, hace la operación mencionada a la derecha de los dos puntos y pasa al estado adonde apunta la flecha. El diagrama indica que cuando la MT en cuestión empieza a funcionar en el estado 0 leyendo el primer palote de una fila de $n$, avanza y se mantiene en el mismo estado. Sigue avanzando en el estado 0 hasta que llega a un blanco. Entonces reemplaza ese blanco con un palote y pasa al estado 1. En ese momento la MT está leyendo el último palote de una fila de $n+1$. La cinta contiene, pues, el resultado buscado. Con todo, según nuestras convenciones, para "recibir" el resultado la MT tiene que estacionarse en el primer cuadrado de la fila que lo representa. Por eso, el programa dispone que la MT retroceda, en el estado 1, hasta encontrar el blanco que abre la fila de palotes que representa el resultado. En cuanto da con él, avanza un cuadrado, con lo cual se coloca ante el primer palote de la fila, y pasa al estado 2. Como el círculo que representa al estado 2 no es el origen de ninguna flecha, se trata de un estado final, y la MT se detiene cuando lo alcanza.

El diagrama (2) representa —como el lector fácilmente comprobará— una

MT que computa la función cero, $x \mapsto 0$. (Recuérdese que representamos el número cero mediante un palote solitario; *vide* p. 387):




(2)

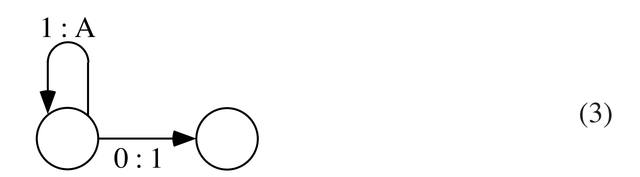
En estos diagramas se pueden discernir componentes que representan tareas parciales dentro de la computación total. Por ejemplo, en el diagrama (1), el componente que forman los estados 0 y 1 con las dos flechas que se originan en el estado 0 simboliza la tarea de buscar un blanco en el cuadrado inicial o a su derecha y escribir un palote en el primero que encuentre; el componente formado por los estados 1 y 2 y las flechas que se originan en el estado 1 simboliza la tarea de buscar un blanco en el cuadrado inicial o a su izquierda y detenerse ante el primer palote a la derecha de ese blanco. Obviamente, tales tareas pueden formar parte de la computación de diversas funciones, en cuyo caso los componentes que representan a aquéllas figurarán en los diagramas de las MT que computan a éstas (con los estados posiblemente numerados de otra manera). Por otra parte, el diagrama de cualquier MT se deja analizar en componentes que representan tareas parciales ejecutadas por esa MT.

Llamo *diagrama parcial* —abreviado DP— a cualquier sistema de círculos unidos con flechas apto para representar del modo descrito una tarea parcial de cómputo, aunque no represente el programa de una MT. Todo DP satisface las condiciones siguientes:

(i) hay un círculo *inicial* que no es el destino de ninguna flecha procedente de otro círculo;
(ii) hay un círculo *final* que no es el origen de ninguna flecha;
(iii) contiene un número finito ($\geq 0$) de círculos intermedios, cada uno de los cuales recibe por lo menos una flecha procedente de otro círculo y origina por lo menos una flecha que apunta a otro círculo;
(iv) cada flecha sale de un círculo y llega a un círculo (posiblemente el mismo), y

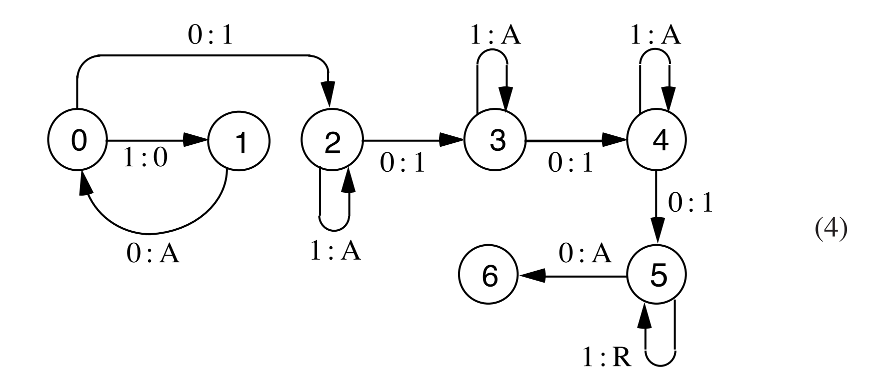
(v) de cada círculo sale *a lo sumo* una flecha marcada 0:$C$ y *a lo sumo* una flecha marcada 1:$C$ (donde $C$ es uno de los caracteres 0, 1, A, R).

Según esta definición, el diagrama de una MT capaz de detenerse es un DP. Si **F** y **G** son dos DP, el DP **F**·**G** se obtiene identificando el círculo final de **F** con el círculo inicial de **G** (esto es, eliminando el círculo final de **F** y llevando todas las flechas que apuntaban a él a desembocar en el círculo inicial de **G**). A veces, digo que **F**·**G** se forma *combinando* **F** con **G** o *acoplando* **G** a **F**. Escribo $\mathbf{F}^2$ en vez de **F**·**F** y $\mathbf{F}^n$ en vez de $\mathbf{F}\cdot\mathbf{F}^{n-1}$ ($n > 1$). Del mismo modo, llamo $(\mathbf{F}\cdot\mathbf{G})^n$ al DP **F**·**G**·**F**·**G**· . . . ·**F**·**G** —formado acoplando $n$ copias de **F**·**G**— el cual, por cierto, no es igual a **F**·**F**· . . . ·**F**·**G**·**G**· . . . ·**G** = $\mathbf{F}^n\cdot\mathbf{G}^n$. Diré generalmente que un DP *hace* o *ejecuta* tal o cual tarea, por decir que la representa. Si $n$ es el número que lleva un círculo en un diagrama dado, el DP $n$ —relativamente a ese diagrama— es el DP formado por ese círculo, con todas las flechas que se originan en él, *suplementado por* un único estado final al que apunten aquellas flechas que no retornan al círculo $n$. Por ejemplo, en el diagrama (1), el DP 0 es éste:



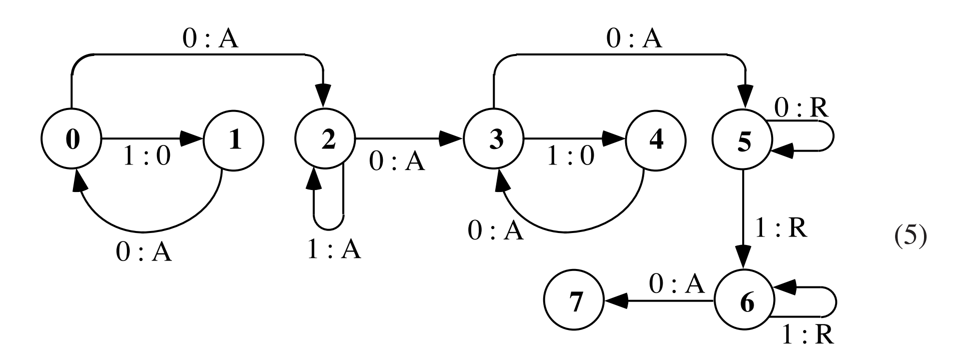

(3)

Apliquemos estos recursos a la descripción de una MT que compute la función constante $x \mapsto n$, esto es, la función que asigna a cualquier número $x$ el sucesor del sucesor . . . del sucesor ($n$ veces) de 0. Evidentemente, su diagrama puede formarse acoplando al diagrama (2) de la función $x \mapsto 0$, $n$ copias (acopladas) del DP (3), y acoplando al DP resultante una copia del DP 1 del diagrama (1), que busca el primer blanco delante suyo o a su izquierda y se detiene ante el primer palote a la derecha de ese blanco (la ejecución de esta última tarea hace que nuestra MT se detenga en la posición prescrita para recibir el resultado que escribió al ejecutar las tareas anteriores). Por ejemplo, el diagrama (4) representa el programa de una MT que computa la función constante $x \mapsto 3$.




(4)

Consideremos ahora la MT que computa la $k$-ésima proyección de $\mathbb{N}^r$, esto es, la función $\langle x_1, \ldots, x_k, \ldots, x_r\rangle \mapsto x_k$. Esta MT debe buscar el ($k$–1)-ésimo blanco a la derecha del cuadrado inicial, borrando a su paso todos los palotes que encuentre en el camino; enseguida debe recorrer la fila de palotes que abre el ($k$–1)-ésimo blanco, preservándola intacta, pues ella representa a $x_k$; luego ha de borrar todos los palotes que encuentre entre el $k$-ésimo y el $r$-ésimo blanco, y finalmente regresar al cuadrado que ocupa el primer palote de la fila preservada y detenerse ahí. Doy el diagrama de la MT que computa $\langle x_1,x_2,x_3\rangle \mapsto x_2$, la segunda proyección de $\mathbb{N}^3$.



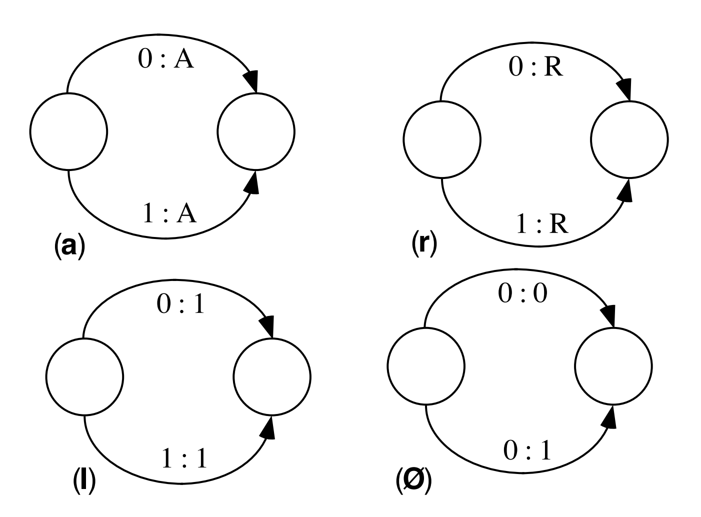

(5)

Sea **F** el DP 0·1 (esto es, el DP que forman los estados 0 y 1). Es fácil ver que el DP 3·4 es una copia de **F**. Advertimos (i) que **F** borra el *número a que se aplica y se detiene ante el primer palote del *número siguiente; (ii) que el DP 2 recorre de comienzo a fin, sin alterarlo, todo *número al que se

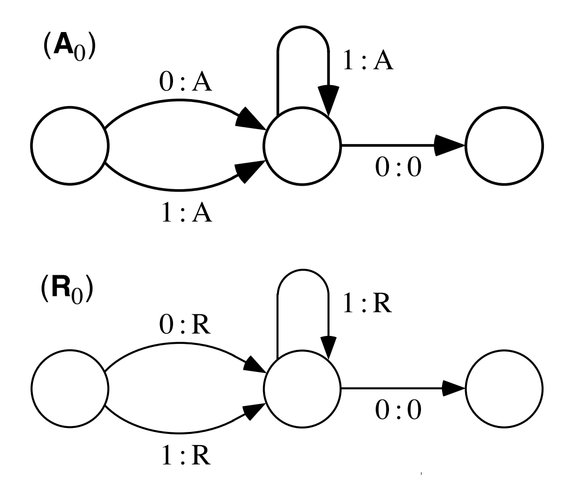
aplica y se detiene en el primer cuadrado después del blanco que cierra ese *número; (iii) que el DP 5 busca el primer palote que haya delante suyo o a su izquierda y se detiene a la izquierda del mismo cuando lo encuentra, y (iv) que el DP 6·7 es igual al DP 5·6 del diagrama (4), cuya tarea es buscar el primer blanco delante suyo o a su derecha y detenerse ante el primer palote a la derecha de ese blanco. Es claro, entonces que el DP $\mathbf{F}^{k-1}\cdot 2\cdot \mathbf{F}^{r-k}\cdot 5\cdot 6\cdot 7$ representa el programa de la MT que computa la *k*-ésima proyección de $\mathbb{N}^r$.

A continuación, enseñaré a construir diversos DP que ejecutan varias tareas básicas de cómputo. Los uso luego (en la Sección 2.11.5) para describir las MT capaces de computar cualquier función recursiva general definida a partir de funciones de esa clase conforme a los esquemas R4, R5 y R6 de la Sección 2.11.1 (o sea, por composición, recursión y búsqueda del número mínimo que cumple cierta condición recursiva). Para facilitar las referencias designaré a cada uno de estos DP con una abreviatura mnemotécnica apropiada.

Los primeros cuatro DP ejecutan las cuatro operaciones elementales de que es capaz una MT: **a** avanza un cuadrado y **r** retrocede un cuadrado, haya lo que haya en el cuadrado inicial; **I** escribe un palote en el cuadrado inicial o lo preserva si ya lo hay, y **Ø** produce un blanco o lo preserva.

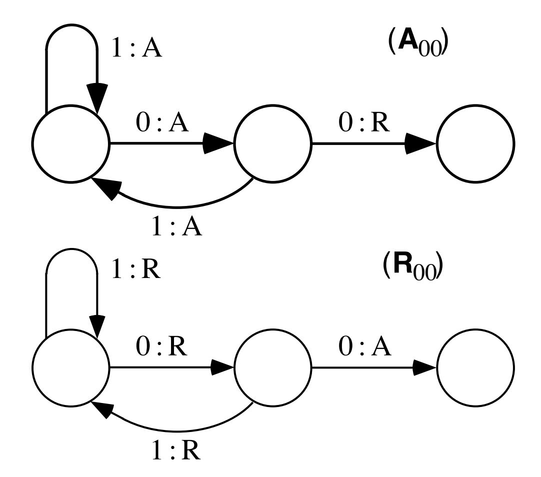



Los dos DP siguientes salen del cuadrado inicial buscando el primer blanco a la derecha ($\mathbf{A}_0$) o a la izquierda ($\mathbf{R}_0$) y se detienen ante él cuando lo encuentran.



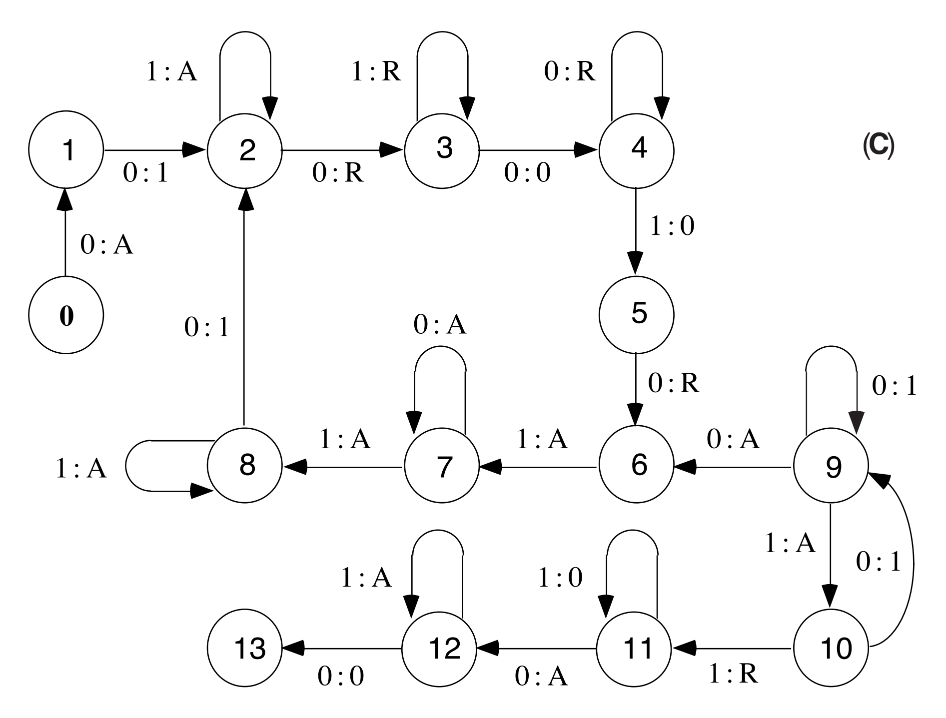

$\mathbf{A}_{|}$ y $\mathbf{R}_{|}$ hacen otro tanto con el primer palote; para dibujarlos, reemplácese simultáneamente 1 por 0 y 0 por 1 en $\mathbf{A}_0$ y $\mathbf{R}_0$.

Un ‘doble blanco’ es un par de blancos consecutivos. $\mathbf{A}_{00}$ busca el primer doble blanco a la derecha del cuadrado inicial y se detiene ante el primero de los dos cuadrados de ese doble blanco. $\mathbf{A}_{00}$ se detiene en el cuadrado inicial si éste forma un doble blanco con el vecino de la derecha. $\mathbf{R}_{00}$ busca el primer doble blanco a la izquierda del cuadrado inicial y se detiene ante el segundo de los cuadrados de ese doble blanco. $\mathbf{R}_{00}$ se detiene en el cuadrado inicial si éste forma un doble blanco con el vecino de la izquierda.




El próximo DP es algo más complejo. La tarea que representa puede des-

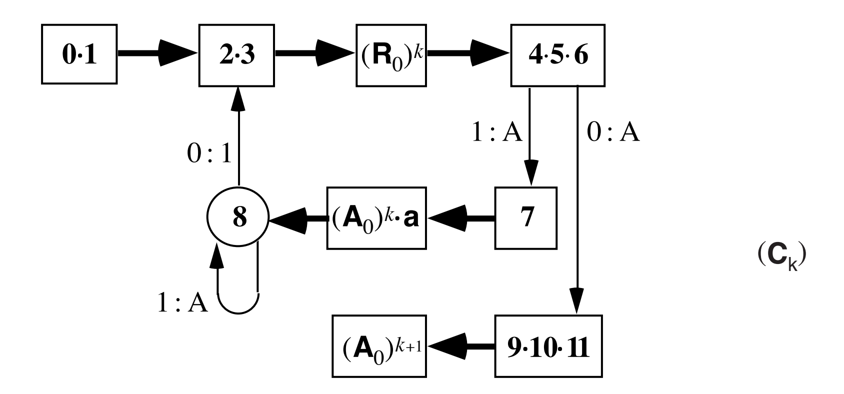
cribirse así: si □ denota el blanco que cierra un cierto *número y no hay ningún palote a la derecha de □, copiar ese *número en la cinta inmediatamente a la derecha de □ y, concluida esa tarea, detenerse en el blanco que cierra la copia. Para facilitar las referencias he numerado los estados.




Para verificar que **C** hace lo que dije, el lector debe escribir tres o cuatro palotes consecutivos en una línea de un papel cuadriculado, situarse en el cuadrado a la derecha del último palote y seguir las instrucciones del diagrama. Si, después de escribir o borrar varios palotes, la línea en que está operando se torna confusa, copie en limpio su última configuración en una línea nueva y siga operando sobre ésta.

Consideremos ahora una variante de la tarea que ejecuta **C**. Supongamos que la cinta contiene un *$n$-tuplo seguido de blancos. Se trata de copiar al final del *$n$-tuplo (inmediatamente después del blanco que lo cierra), el ($n$–$k$)-ésimo *número del mismo y detenerse en el blanco que cierra la copia. Llamaré $\mathbf{C}_k$ al DP que ejecuta esta tarea. (Con esta nueva nomenclatura, $\mathbf{C} = \mathbf{C}_0$). Supondremos que $\mathbf{C}_k$ parte del blanco □ que cierra el *$n$-tuplo.

Contando de izquierda a derecha, sea $_k\square$ el $k$-ésimo blanco anterior a $\square$. $_k\square$ es el blanco que cierra el *número que $\mathbf{C}_k$ va a copiar. $\mathbf{C}_k$ tiene que cruzar varias veces de izquierda a derecha y de derecha a izquierda, sin alterarlo, el *$k$-tuplo situado entre $_k\square$ y $\square$. Para asegurar que esto ocurra hay que insertar en puntos apropiados de **C** una cantidad suficiente de copias de los DP $\mathbf{R}_0$ (que busca el primer blanco a la izquierda de la posición ocupada) y $\mathbf{A}_0$ (que busca el primer blanco a la derecha de la posición ocupada). Igual que **C**, $\mathbf{C}_k$ empieza escribiendo un palote a la derecha de $\square$ y retrocediendo a leer el *número que debe copiar. Entre 3 y 4 la copiadora retrocede de $\square$ a $\square_k$, esto es, ejecuta la tarea del DP $(\mathbf{R}_0)^k$. Entre 7 y 8 va del primer palote del *$k$-tuplo al primer palote de la copia, para lo cual tiene que llegar a $\square_k$ y avanzar un cuadrado: esa es la tarea de $(\mathbf{R}_0)^k\cdot\mathbf{a}$. Por último, cuando ha terminado de hacer la copia y, después de restaurar en su sitio original el *número que copió, llega a $_k\square$ en estado 11, $\mathbf{C}_k$ repite la travesía de $_k\square$ a $\square$ y sigue hasta el primer blanco a la derecha de $\square$, que es el blanco que cierra la copia; en otras palabras, ejecuta la tarea del DP $(\mathbf{A}_0)^{k+1}$. A continuación represento gráficamente este análisis mediante un diagrama algo diferente de los utilizados arriba.



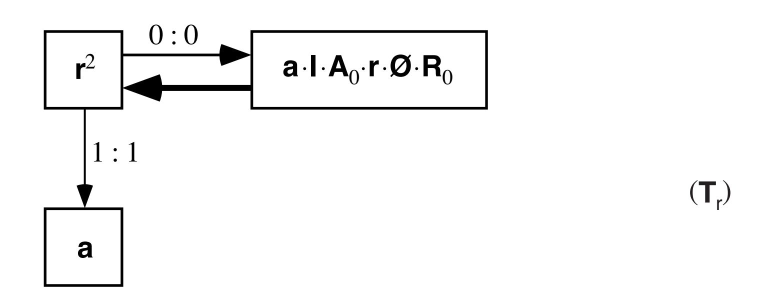

Cada rectángulo representa el DP cuyo nombre contiene. Se entiende que una flecha que une dos rectángulos $A$ y $B$ va del círculo final del DP representado por $A$ al círculo inicial del DP representado por $B$. Una flecha gruesa significa que la transición indicada ocurre cualquiera que sea el signo leído y envuelve solamente un cambio de estado (en otras palabras, una flecha gruesa representa dos flechas corrientes, con el mismo origen y el mismo destino, acompañadas, respectivamente, de las indicaciones 0:0 y 1:1).

En el curso de una computación puede ser necesario desplazar un *número en una dirección en que la cinta está completamente en blanco. Consideremos la forma más simple de esta tarea. Llamo $\mathbf{\Delta}_r$ al DP que empieza a operar inmediatamente a la izquierda de un *número que sólo está precedido por blancos, desplaza ese *número un cuadrado hacia la izquierda y se detiene en el blanco que precede al *número desplazado. Obviamente

$$\mathbf{\Delta}_r = \mathbf{I}\cdot\mathbf{A}_0\cdot\mathbf{Ø}\cdot\mathbf{R}_{00}$$

$\mathbf{\Delta}_a$ ejecuta la tarea opuesta; para describirlo reemplácese 'izquierda' por 'derecha' y 'precede' por 'sigue' —o intercámbiense **A** y **R**— en la ecuación precedente.

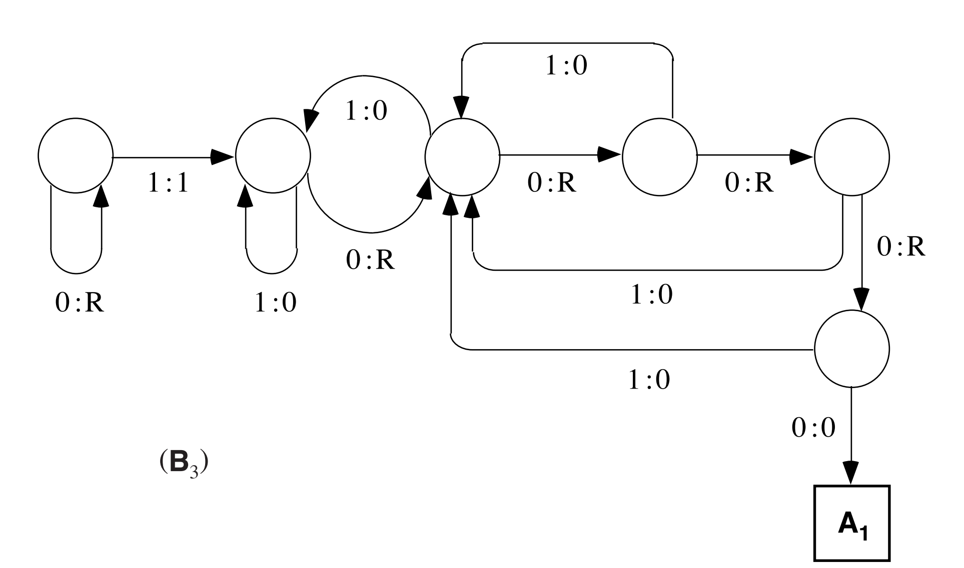
También puede ser necesario acercar dos *números que estén separados en la cinta por más de un blanco. El DP siguiente, que llamaré $\mathbf{T}_r$, empieza a operar en el blanco que precede inmediatamente a un *número y traslada este *número hacia atrás hasta que entre él y el *número anterior no queda más que un blanco, en el cual $\mathbf{T}_r$ se detiene.




($\mathbf{T}_r$)

Obsérvese cómo $\mathbf{T}_r$ retrocede inicialmente un cuadrado; si encuentra un palote, hay un solo blanco entre él y el *número siguiente; por ende, $\mathbf{T}_r$ no tiene nada que hacer; avanza a ese blanco y se detiene. Pero si encuentra un blanco, retorna al blanco inicial, lo llena con un palote, avanza al extremo derecho del *número que va a trasladar, le quita el último palote, busca el blanco que *ahora* precede ese *número y recomienza su trabajo. El lector puede entretenerse en diseñar $\mathbf{T}_a$, que hace el trabajo inverso: empieza en el blanco que sigue inmediatamente a un *número y traslada ese *número hacia adelante hasta que entre él y el *número siguiente no queda más que un blanco, en el cual $\mathbf{T}_a$ se detiene.

Es cómodo programar la MT que computa una función $f$ de tal modo que escriba el valor de $f$ a la derecha del argumento y de cualesquiera *números auxiliares que tenga que escribir mientras saca cuentas. Para que "reciba" el resultado, conforme a nuestra descripción canónica, estacionada en el primer palote de la fila que lo representa, en una cinta por lo demás en blanco, la MT tendrá entonces que borrar primero los palotes que representan el argumento y todos los *números auxiliares. El DP $\mathbf{B}_n$ hace eso precisamente, si la cinta no contiene ningún *número separado por más de $n$ blancos del *número más próximo a su derecha: partiendo del blanco que precede inmediatamente al resultado, $\mathbf{B}_n$ retrocede borrando todos los palotes a su izquierda. Cuando $\mathbf{B}_n$ ha borrado el último palote en esa dirección, retrocede $n$+1 cuadrados y al detectar un blanco se pone a buscar el primer palote a su derecha —que será justamente el primer palote del resultado— y se detiene allí. Si todo lo que hay que borrar es el argumento de $f$, se puede tomar $n$ = 1, que es la separación entre los *números de un *$r$-tuplo. Pero puede ocurrir que el programa de la MT en cuestión prescriba anotar *números auxiliares con una separación mayor (para distinguirlos). En todo caso, $n$ será finito. Doy el diagrama de $\mathbf{B}_3$, que tendremos ocasión de utilizar en la Sección 2.11.5.



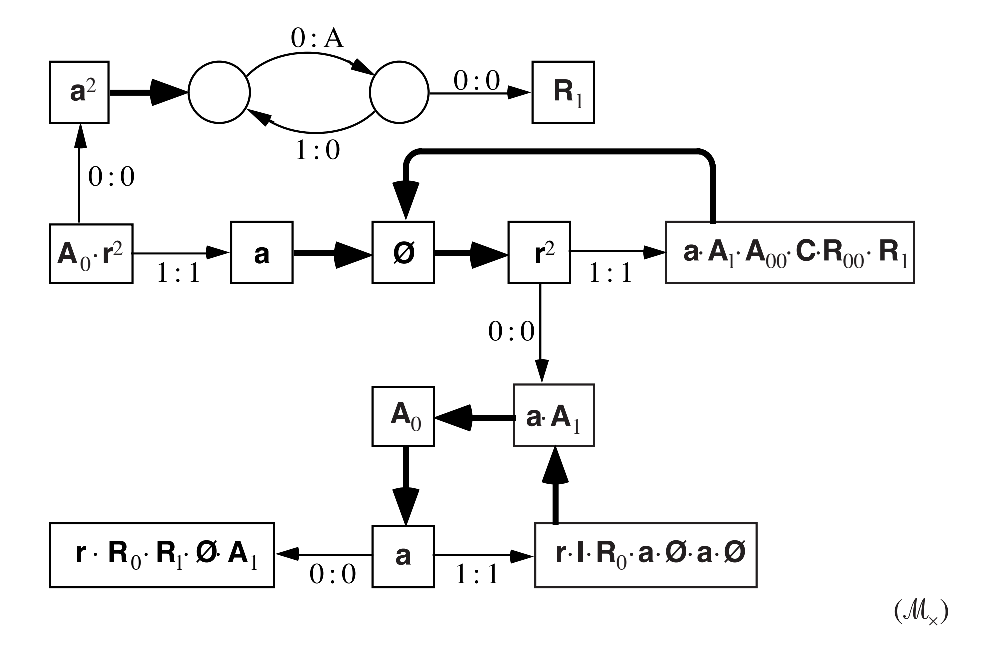

($\mathbf{B}_3$)

Daré dos ejemplos de MT que computan funciones aritméticas ordinarias. $\mathcal{M}_+$ computa la función $\langle x,y\rangle \mapsto x + y$, esto es, la suma de cualquier par de números naturales.

$$\mathcal{M}_+ = \mathbf{A}_0 \cdot \mathbf{I} \cdot \mathbf{A}_0 \cdot (\mathbf{r} \cdot \text{Ø})^2 \cdot \mathbf{R}_0 \cdot \mathbf{a}.$$

Dada una fila de $x$+1 palotes seguida de un blanco seguido de una fila de $y$+1 palotes, si $\mathcal{M}_+$ empieza a operar, conforme a la descripción canónica, en el primer cuadrado de la primera fila, busca el primer blanco a la derecha, esto es, el que separa las dos filas de palotes; lo llena con un palote; busca nuevamente el primer blanco a la derecha; retrocede dos veces, borrando al paso dos palotes; busca el primer blanco a la izquierda y avanza un cuadrado, con lo cual acaba situada ante el primer palote de una fila de $x + y + 1$ palotes consecutivos, que es justamente el *número que representa a $x + y$.

$\mathcal{M}_\times$ computa la función $\langle x,y\rangle \mapsto xy$, esto es, el producto de cualquier par de números naturales, El programa de $\mathcal{M}_\times$ es, por cierto, bastante más complejo que el de $\mathcal{M}_+$. Lo representaré mediante un diagrama como el que usé para $\mathbf{C}_k$.




($\mathcal{M}_\times$)

Veamos cómo $\mathcal{M}_\times$ computa el producto $xy$. Sean $*x$ e $*y$ los *números que representan en la cinta a nuestros dos factores. $\mathcal{M}_\times$ avanza desde su posición canónica inicial sobre el primer palote de $*x$ hasta el blanco $\square$ entre $*x$ e $*y$. Retrocede dos cuadrados. Si $x = 0$, $\mathcal{M}_\times$ encuentra un blanco. Su tarea restante —ejecutada por el DP en la línea superior del diagrama— consiste en borrar $*y$ e ir a pararse sobre el palote único de $*x$. Pero si el segundo cuadrado anterior a $\square$ no está en blanco, $x \geq 1$ y la tarea restante de $\mathcal{M}_\times$ consiste en (i) escribir a la derecha de $*y$ un $*(x–1)$-tuplo de copias de $*y$, para que haya en la cinta $x$ representaciones de $y$; (ii) llenar con palotes los blancos que separan esas $x$ representaciones de $y$, para que formen una fila consecutiva de palotes; (iii) borrar los palotes que esa fila tiene en exceso de $xy + 1$, el número de palotes de $*(xy)$; (iv) borrar $*x$, y (v) ir a detenerse sobre el primer palote de $*(xy)$. El lector debe comprobar que el DP cuyo componente inicial (marcado con **a**) está a la derecha del componente inicial de $\mathcal{M}_\times$ ejecuta precisamente esta conjunción de tareas.

De acuerdo con nuestra descripción canónica, la MT que computa una función numérica recibe una cinta en que sólo está representado un argumento y entrega una cinta en que sólo está representado el valor correspondiente. La MT borra, pues, los datos antes de entregar el resultado. Esta convención es razonable, pero resulta incómoda en ciertas ocasiones. Por ejemplo, para computar la función compuesta

$$\langle x_1, \dots ,x_n\rangle \mapsto h(g_1(x_1, \dots ,x_n), \dots , g_m(x_1, \dots ,x_n))$$

donde $h$ es una función T-computable $m$-aria y $g_1, \dots , g_m$ son funciones T-computables $n$-arias, conviene disponer todo el tiempo del dato $\langle x_1, \dots ,x_n\rangle$ mientras se computan sucesivamente los valores respectivos de $g_1, \dots , g_m$, y tener todos estos valores a mano para computar el valor buscado de $h$. Con este propósito, asociaré a *cada* función T-computable $f$ un diagrama parcial $\mathbf{D}_f$ que entrega el valor de $f$ para cada argumento que se le suministre, sin borrar de la cinta la representación de ese argumento. $\mathbf{D}_f$ empieza a trabajar en el blanco que cierra la representación de $\langle x_1, \dots ,x_n\rangle$, escribe a su derecha una fila de $f(x_1, \dots ,x_n)+1$ palotes y se detiene en el blanco que cierra esta fila.

No es difícil probar que una función $n$-aria $f$ puede asociarse con un DP $\mathbf{D}_f$ que cumpla las condiciones señaladas si y solo si existe una MT $\mathcal{M}_f$ que

computa a $f$ con arreglo a nuestra descripción canónica.[38] Llamaré $\mathbf{M}_f$ al DP que ejecuta la tarea de $\mathcal{M}_f$. (La diferencia más significativa entre $\mathbf{M}_f$ y $\mathcal{M}_f$ consiste en que $\mathbf{M}_f$ puede trabajar en una cinta que contenga palotes a la izquierda del *$n$-tuplo a que se aplica, cosa que, según nuestra descripción canónica, $\mathcal{M}_f$ no puede hacer). Si $\mathbf{D}_f$ está dado, podemos poner

$$\mathbf{M}_f = \mathbf{r}_{00} \cdot \mathbf{D}_f \cdot \mathbf{R}_0 \cdot \mathbf{B}_1$$

$\mathbf{M}_f$ empieza en el primer palote del *$n$-tuplo que representa el argumento, busca el primer doble blanco a su derecha, ejecuta la tarea de $\mathbf{D}_f$, retrocede al blanco que separa el argumento del correspondiente valor, y finalmente ejecuta la tarea de $\mathbf{B}_1$ que borra el argumento y se detiene en el primer palote del valor. Supongamos ahora que $\mathbf{M}_f$ está dada y que sólo incluye "borradores" del tipo $\mathbf{B}_k$ con $k \leq q$. Ponemos entonces

$$\mathbf{D}_f = \mathbf{a} \cdot \mathbf{l} \cdot \mathbf{a} \cdot (\mathbf{C}_\mathrm{n})^{n} \cdot (\mathbf{R}_0)^{n} \cdot \mathbf{r} \cdot \mathbf{Ø} \cdot \mathbf{R}_\mathrm{l} \cdot \mathbf{R}_{00} \cdot ((\mathbf{\Delta}_\mathrm{r})^{q} \cdot \mathbf{A}_0 \cdot \mathbf{A}_\mathrm{l} \cdot \mathbf{r})^{n} \cdot \mathbf{a} \cdot \mathbf{M}_f \cdot \mathbf{r} \cdot \mathbf{T}_\mathrm{r} \cdot \mathbf{A}_{00}$$

Analicemos las tareas que $\mathbf{D}_f$ cumple antes de ejecutar la tarea de $\mathbf{M}_f$. Empieza a operar en el blanco a la derecha del *$n$-tuplo que representa el argumento. Avanza un cuadrado, escribe un palote y avanza otro cuadrado, situándose en un blanco que llamaré ☐. La tarea siguiente, ejecutada por el DP $(\mathbf{C}_\mathrm{n})^n$, consiste en copiar el argumento a la derecha de ☐. (Recuérdese que $\mathbf{C}_\mathrm{n}$ copia a la derecha del blanco inicial el *número que está separado por $n$ *números de ese blanco y se detiene en el blanco que cierra la copia; en su primera aplicación, pues, $\mathbf{C}_\mathrm{n}$ copia el primer *número del argumento, que está separado de ☐ por los $n$–1 *números restantes y el palote recién escrito; en la segunda aplicación, $\mathbf{C}_\mathrm{n}$ copia el segundo *número del argumento, . . . , en la $n$-ésima, el $n$-ésimo). Luego, el DP retorna a ☐ (que ahora es el $n$-ésimo blanco a la derecha del blanco donde se concluye la última aplicación de $\mathbf{C}_\mathrm{n}$), retrocede un cuadrado, borra el palote que escribió al

38 En verdad, $\mathbf{D}_f$ no es sino una máquina de Turing que computa $f$ de acuerdo con una descripción canónica diferente, que pude muy bien adoptar —como hace Hermes (1961)— en vez de la que dí en la Sección 2.11.3; pero esta última tiene ventajas didácticas.

principio, y busca el doble blanco que precede al *$n$-tuplo original. La tarea siguiente, ejecutada por el DP $((\mathbf{\Delta}_\mathrm{r})^{q}\cdot\mathbf{A}_0\cdot\mathbf{A}_\mathrm{l}\cdot\mathbf{r})^n$, consiste en desplazar el *$n$-tuplo original $q$ cuadrados a la izquierda, para ponerlo fuera del alcance de los "borradores" incluidos en $\mathbf{M}_f$. En efecto, $(\mathbf{\Delta}_\mathrm{r})^q$ desplaza $q$ cuadrados a la izquierda el *número que esté inmediatamente a la derecha del blanco donde empieza a operar; $\mathbf{A}_0\cdot\mathbf{A}_\mathrm{l}$ llevan al primer palote del *número siguiente y $\mathbf{r}$ al blanco que inmediatamente lo precede. La $n$-ésima aplicación de $((\mathbf{\Delta}_\mathrm{r})^{q}\cdot\mathbf{A}_0\cdot\mathbf{A}_\mathrm{l}\cdot\mathbf{r}$ desplaza $q$ cuadrados a la izquierda al último *número del *$n$-tuplo referido y se detiene en el blanco inmediatamente a la derecha *de la copia*; la copia estará separada entonces por $q$+3 blancos del último *número del *$n$-tuplo original.[39] Avanzando un cuadrado, nos colocamos sobre el primer palote de la copia, a la que se aplica $\mathbf{M}_f$. Como sabemos, la tarea de $\mathbf{M}_f$ consiste en reemplazar esa copia, por el *número que representa el valor correspondiente de $f$, eliminando con sus "borradores" todo *número que preceda a ese valor y diste a lo sumo $q$ blancos del *número siguiente. Conforme a la descripción canónica, $\mathbf{M}_f$ se detiene en el primer palote del resultado. A gran distancia a la izquierda la cinta contiene el *$n$-tuplo original. La tarea final de $\mathbf{D}_f$, ejecutada por $\mathbf{r}\cdot\mathbf{T}_\mathrm{r}$, consiste en trasladar a la izquierda el resultado hasta que entre éste y el *$n$-tuplo original no haya más que un solo blanco, y avanzar desde allí hasta llegar al blanco que cierra el resultado, donde $\mathbf{D}_f$ se detiene.

39 Para poner al *$n$-tuplo original fuera del alcance de los "borradores" de $\mathbf{M}_f$ basta en rigor con una separación de $q$+1 blancos, de modo que sería suficiente utilizar $(\mathbf{\Delta}_\mathrm{r})^{q-2}$ en vez de $(\mathbf{\Delta}_\mathrm{r})^q$.

### *2.11.5 Demostración de resultados*

*I. Toda función recursiva es T-computable*. Hay que probar que cualquier función numérica definida según uno de los esquemas R1-R6 de la Sección 2.11.1 es T-computable.[40] Los diagramas (1) y (2) de la Sección 2.11.3 exhiben las MT que computan la función sucesor $\sigma{:}x \mapsto x + 1$ y la función cero $x \mapsto 0$, respectivamente. El diagrama (5) exhibe, a título de ejemplo, la MT que computa $\langle x_1,x_2,x_3\rangle \mapsto x_2$, la segunda proyección de $\mathbb{N}^3$. En general, la $k$-ésima proyección de $\mathbb{N}^r$ es computada por $\mathbf{A}_{00}{\cdot}\mathbf{C}_{r-k}{\cdot}\mathbf{A}_0{\cdot}\mathbf{B}_1$.

Si $h$ es una función recursiva $m$-aria T-computable, asociada al DP $\mathbf{D}_h$ y $g_1, \ldots ,g_m$ son funciones recursivas $n$-arias T-computables, asociadas a los DP $\mathbf{D}_1, \ldots , \mathbf{D}_m$, respectivamente, la función compuesta $f{:}\langle x_1, \ldots ,x_n\rangle \mapsto h(g_1(x_1, \ldots ,x_n), \ldots ,g_m(x_1, \ldots ,x_n))$ es recursiva (por R4) y T-computable por la MT descrita a continuación:

$$\mathcal{M}_f = \mathbf{A}_{00}{\cdot}\mathbf{a}{\cdot}\mathbf{l}{\cdot}\mathbf{a}{\cdot}(\mathbf{C}_n)^{n}{\cdot}\mathbf{R}^{n}{\cdot}\mathbf{r}{\cdot}\boldsymbol{\varnothing}{\cdot}\mathbf{A}_{00}{\cdot}\mathbf{D}_1{\cdot}(\mathbf{C}_n)^{n}{\cdot}\mathbf{D}_2{\cdot}\ldots{\cdot}(\mathbf{C}_n)^{n}{\cdot}\mathbf{D}_m{\cdot}$$
$$\mathbf{C}_{(m-1)+n(m-1)}{\cdot}\mathbf{C}_{(m-1)+n(m-2)}{\cdot}\ldots{\cdot}\mathbf{C}_{m-1}{\cdot}\mathbf{D}_h{\cdot}\mathbf{R}_0{\cdot}\mathbf{B}_3$$

El funcionamiento de $\mathcal{M}_f$ es tan sencillo como ingenioso. Después de situarse en el primer blanco a la derecha del *$n$-tuplo que representa el argumento $\langle x_1, \ldots ,x_n\rangle$, escribe un palote en el blanco siguiente, avanza al próximo blanco y copia ese *$n$-tuplo; luego borra el palote auxiliar, dejando tres blancos entre el original y la copia; avanza al blanco que cierra la copia y escribe a su derecha el *número que representa a $g_1(x_1, \ldots ,x_n)$; $(\mathbf{C}_n)^n$ produce una segunda copia del argumento a la derecha de ese *número y $\mathbf{D}_2$ escribe el *número que representa a $g_2(x_1, \ldots ,x_n)$ a la derecha del blanco que cierra esa segunda copia. Los puntos suspensivos indican la repetición de $(\mathbf{C}_n)^{n}{\cdot}\mathbf{D}_k$ con $k = 3, \ldots , n-1$. Cuando $\mathbf{D}_m$ concluye su trabajo, $\mathcal{M}_f$ se encuentra situada en el blanco que cierra un *$m(n+1)$-tuplo formado por representaciones de $\langle x_1, \ldots ,x_n,g_k(x_1, \ldots ,x_n)\rangle$ $(1 \leq k \leq m)$, con un blanco intercalado entre cada dos representaciones consecutivas. La tarea siguiente, efectuada por el DP $\mathbf{C}_{(m-1)+n(m-1)}{\cdot}\mathbf{C}_{(m-1)+n(m-2)}{\cdot}\ldots{\cdot}\mathbf{C}_{m-1}$ consiste en copiar, a continuación del

[40] Con R6 ajustado a las indicaciones del párrafo que le sigue.

*$m(n+1)$-tuplo indicado, precisamente los $m$ *números que representan a $g_1(x_1, \ldots, x_n), \ldots, g_m(x_1, \ldots, x_n)$.[41] Por último, **D**$_h$, operando desde el blanco que cierra este *$m$-tuplo, escribe a su derecha el *número buscado, que representa a $h(g_1(x_1, \ldots, x_n), \ldots, g_m(x_1, \ldots, x_n)) = f(x_1, \ldots, x_n)$. **R**$_0$·**B**$_3$ da los últimos toques necesarios para entregar el resultado en la forma convencional prescrita en nuestra descripción canónica.

Sea $g_1$ una función recursiva $n$-aria T-computable asociada al DP **D**$_1$ (o una constante, si $n = 0$) y $g_2$ una función recursiva $(n+2)$-aria T-computable asociada al DP **D**$_2$. Sea $f$ la función $(n+1)$-aria definida por recursión así:[42]

$$f(x_1, \ldots, x_n, 0) = g_1(x_1, \ldots, x_n)$$
$$f(x_1, \ldots, x_n, y+1) = g_2(x_1, \ldots, x_n, y, f(x_1, \ldots, x_n, y))$$

41 Comprobemos este aserto, trabajando, por ejemplo, con $m = 3$ y $n = 2$. Llamo $\mathsf{x}$ al *número que representa el número $x$, $\mathsf{g}_1$ a la representación de $g_1(x_1, x_2)$. $\square$ es un blanco. Empezamos a la derecha de la fila $\mathsf{x}_1\square\mathsf{x}_2\square\mathsf{g}_1\square\mathsf{x}_1\square\mathsf{x}_2\square\mathsf{g}_2\square\mathsf{x}_1\square\mathsf{x}_2\square\mathsf{g}_3$. **C**$_6$ copia $\mathsf{g}_1$. Luego **C**$_4$, empezando a la derecha de $\mathsf{x}_1\square\mathsf{x}_2\square\mathsf{g}_1\square\mathsf{x}_1\square\mathsf{x}_2\square\mathsf{g}_2\square\mathsf{x}_1\square\mathsf{x}_2\square\mathsf{g}_3\square\mathsf{g}_1$, copia $\mathsf{g}_2$. Luego, empezando a la derecha de $\mathsf{x}_1\square\mathsf{x}_2\square\mathsf{g}_1\square\mathsf{x}_1\square\mathsf{x}_2\square\mathsf{g}_2\square\mathsf{x}_1\square\mathsf{x}_2\square\mathsf{g}_3\square\mathsf{g}_1$, **C**$_2$ copia $\mathsf{g}_3$.

42 Mi definición de $f$ se aparta un poco del esquema R5, con el propósito de hacer más sencillo el diagrama de la MT que la computa. Para tranquilizar al lector desconfiado, mostraré aquí que la diferencia es inocua. Supongamos, pues, que $\varphi$ es la función $(n+1)$-aria definida, según el esquema R5, por las relaciones

$$\varphi(0, x_1, \ldots, x_n) = \gamma_1(x_1, \ldots, x_n)$$
$$\varphi(y+1, x_1, \ldots, x_n) = \gamma_2(y, \varphi(x_1, \ldots, x_n, y), x_1, \ldots, x_n)$$

donde $\gamma_1$ y $\gamma_2$ son funciones recursivas y T-computables y las funciones $g_1$ y $g_2$ satisfacen las condiciones $g_1 \equiv \gamma_1$ y $g_2(x_1, \ldots, x_{n+2}) = \gamma_2(x_{n+1}, x_{n+2}, x_1, \ldots, x_n)$. Mostraré que $g_2$ es recursiva y T-computable. Sea $\pi_k$ la $k$-ésima proyección de $\mathbb{N}^{n+2}$; sabemos que $\pi_k$ es recursiva (por R3) y T-computable. Por consiguiente, también lo es la función compuesta

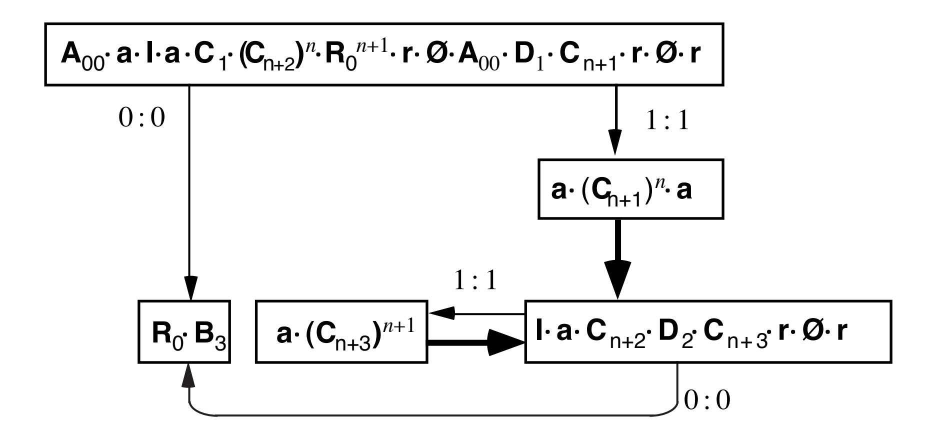
$$g_2: \langle x_1, \ldots, x_{n+2}\rangle \mapsto \gamma_2(\pi_{n+1}(x_1, \ldots, x_{n+2}), \pi_{n+2}(x_1, \ldots, x_{n+2}), \pi_1(x_1, \ldots, x_{n+2}), \ldots, \pi_n(x_1, \ldots, x_{n+2}))$$

Por lo tanto, la función $f$ definida arriba en el texto es T-computable por la MT ahí descrita. Es claro, entonces, que la función compuesta

$$\varphi: \langle x_1, \ldots, x_{n+1}\rangle \mapsto f(\pi_{n+1}(x_1, \ldots, x_{n+1}), \pi_1(x_1, \ldots, x_{n+1}), \ldots, \pi_n(x_1, \ldots, x_{n+1}))$$

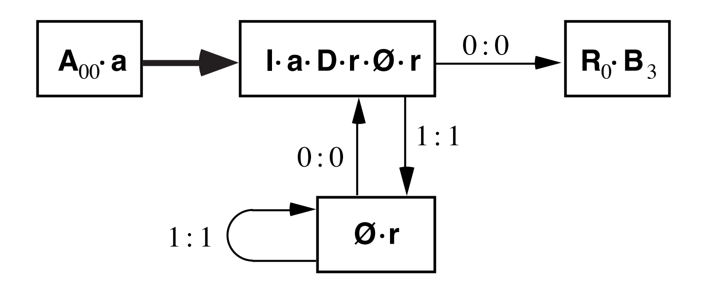
también es T-computable.

$f$ es recursiva (por R5) y T-computable por la MT representada en el diagrama siguiente:




Dejo a cargo del lector el análisis de su funcionamiento (para simplificar, suponga que $n$ = 1).

Sea $g$ una función recursiva general ($n$+1)-aria T-computable, asociada al DP **D** ($n \geq 0$). Digamos que $f$ es la función $n$-aria definida por $f(x_1, \ldots, x_n) = \mu y(g(x_1, \ldots, x_n, y) = 0$. Entonces $f$ es recursiva (por R6) y T-computable por la MT representada en el diagrama siguiente:




¿Cómo trabaja? Supongamos que se aplica al *$n$-tuplo que representa el argumento $\langle x_1, \ldots, x_n \rangle$. Partiendo de su primer palote, busca el blanco que lo cierra, avanza un cuadrado, escribe un palote, avanza otro cuadrado y aplica **D** al *($n$+1)-tuplo precedente, que representa $\langle x_1, \ldots, x_n, 0 \rangle$. Si **D** produce un solo palote, quiere decir que $g(x_1, \ldots, x_n, 0) = 0$. La ejecución de **r·Ø·r** colo-

ca entonces a nuestra MT ante un blanco, inmediatamente a la derecha del resultado buscado. Como siempre, $\mathbf{R}_0 \cdot \mathbf{B}_3$ asegura que éste se entregue en la forma convenida. Pero si **D** produce más de un palote, $g(x_1, \ldots ,x_n,0) \neq 0$. La MT borra entonces el número que escribió **D**, agrega un palote al último *número del *($n$+1)-tuplo precedente, y repite el ejercicio anterior. De esto modo, computará sucesivamente $g(x_1, \ldots ,x_n,1)$, $g(x_1, \ldots ,x_n,2)$, . . . hasta dar con el primer número $h$ tal que $g(x_1, \ldots ,x_n,h) = 0$.

Así queda demostrado que cualquier función recursiva es T-computable.

*II. Toda función T-computable es recursiva*. Sea $f$ una función $r$-aria T-computable por la MT $\mathcal{M}_f$. Ya sabemos identificar a $\mathcal{M}_f$ mediante un gödel. Mientras $\mathcal{M}_f$ computa el valor de $f$ para un dado argumento $x \in \mathbb{N}^r$, la cinta presenta sucesivamente distintas configuraciones; cada momento o etapa de la computación queda exhaustivamente descrito por la indicación de (i) la configuración de la cinta en esa etapa; (ii) el estado momentáneo de $\mathcal{M}_f$ y (iii) el cuadrado que $\mathcal{M}_f$ está leyendo. Se puede definir un gödel que represente toda esta información. Como el tránsito de cada etapa a la próxima está determinado por el programa finito de $\mathcal{M}_f$, se pueden definir funciones recursivas que asignen al gödel de cada etapa, el gödel de la etapa siguiente. Sobre esta base, es posible definir una función recursiva que asigna el valor $f(x)$ al par formado por el gödel de $\mathcal{M}_f$ y el gödel de la etapa inicial de la computación de $f(x)$. Tal es el método estándar para demostrar este resultado,[43] que produce al mismo tiempo pruebas de los resultados III-VI. Pero aquí seguiré otro método más simple, debido a Boolos y Jeffrey (1980, Cap. 8); luego completaré esquemáticamente la demostración estándar a propósito del resultado IV.

Para llevar a cabo la demostración propuesta necesitaré algunos conceptos auxiliares, que explico a continuación en párrafos numerados del 1 al 6.

*1. Definición de una función mediante una lista finita de condiciones*. Sea $\varphi$ una función numérica definida en $\mathcal{D} \subseteq \mathbb{N}^r$ mediante $n$ condiciones de la forma

$$\varphi(x) = \psi_k(x) \text{ si y sólo si el } r\text{-tuplo } x \in C_k \; (1 \leq k \leq n)$$

43 Ideado por Kleene; véase su tratado didáctico (1952), o la exposición muy clara y precisa de Davis (1958, Capítulo 4).

donde $\{C_1, \ldots, C_n\}$ es una partición de $\mathcal{D}$. Sea $\chi_k$ la función característica de $C_k$. Obviamente, $\varphi$ es la función compuesta definida por

$$\varphi(x) = \psi_1(x)(1 - \chi_1(x)) + \ldots + \psi_n(x)(1 - \chi_n(x))$$

Decimos que $\varphi$ es una función definida mediante una lista finita de condiciones. Si $\psi_1, \ldots, \psi_n, \chi_1, \ldots, \chi_n$ son funciones pr-recursivas, $\varphi$ también es pr-recursiva.

*2. Maximización acotada*. El lector conoce la función $r$-aria definida por "minimización acotada" $x \mapsto \mu y(y \leq \phi(x) \wedge R(y,x))$, cuyo valor, para cada $r$-tuplo $x$, es el mínimo número $y \leq \phi(x)$ que cumple la condición $R(y,x)$. Como indiqué en el paso 3º de la prueba del primer teorema de incompletud de Gödel (p. 334, (✣)), esta función es pr-recursiva si la función $\phi$ y la relación $R$ lo son. Ahora definiré por *maximización acotada* la función $x \mapsto \mathrm{M}y(y \leq \phi(x) \wedge R(y,x))$, cuyo valor es el *máximo* número $y \leq \phi(x)$ que cumple la condición $R(y,x)$. Sea $P$ la condición definida por $P(y,x) \leftrightarrow (R(y,x) \wedge \forall z(z \leq \phi(x) \rightarrow (z > y \rightarrow \neg R(z,x))$. Obviamente, $\mathrm{M}y(y \leq \phi(x) \wedge R(y,x)) = \mu y(y \leq \phi(x) \wedge P(y,x))$. Por lo tanto, nuestra función de maximización acotada es pr-recursiva si la función $\phi$ y la condición $R$ lo son. En particular, defino por maximización acotada la función $\theta$ que asigna a cada $x \in \mathbb{N}$ el máximo número $w \leq x$ tal que $2^w \leq x$; en el acto advertimos que $\theta(2^{z+1}-1) = z$.

*3. Codificación de r-tuplos*. Sea $p_k$ el $k$-ésimo número primo. La función $r$-aria

$$\Gamma_r{:}\langle x_1,x_2, \ldots ,x_r\rangle \mapsto 2^{x_1}{\cdot}3^{x_1}{\cdot} \ldots {\cdot}p_r^{x_r}$$

es pr-recursiva (cf. pp. 336 y ss.) y puede utilizarse para identificar inequívocamente un $r$-tuplo numérico mediante un solo número (un gödel). En vez de $\Gamma_r(x_1,x_2, \ldots ,x_r)$ escribimos $[x_1,x_2, \ldots ,x_r]$. Si $x$ designa el $r$-tuplo $\langle x_1,x_2, \ldots ,x_r\rangle$, escribo $[x]$ por $\Gamma_r(x)$. Consideremos ahora la función $\Pi_k$ que asigna a cada número $x$ el máximo número $z \leq x$ tal que $x$ es divisible por $(p_k)^z$. Según lo explicado en el párrafo 2, $\Pi_k$ es pr-recursiva. Obsérvese que, si $x$ es el gödel de un $r$-tuplo y $1 \leq k \leq r$, $\Pi_k(x)$ es el $k$-ésimo número del $r$-tuplo en cuestión. En otras palabras $\Pi_k([x]) = \pi_k(x)$.

*4. Codificación de computaciones*. Sea $\mathcal{M}_f$ una MT que computa una función $r$-aria $f$. Las etapas de la computación de $f(x)$ pueden numerarse, diga-

mos, de 0 a $n$. En la etapa $t$, $\mathcal{M}_f$ lee un determinado cuadrado de su cinta, escrito o en blanco, que llamaré $\square^t$. Si equiparamos cada palote con el dígito 1 y cada blanco con el dígito 0, podemos leer el contenido de la cinta, a la izquierda de $\square^t$, como la representación de un número natural en el sistema binario. Llamo a ese número, el *número de la izquierda* $\rho_t$ correspondiente a la etapa $t$. El contenido del resto de la cinta, a partir de $\square^t$, también puede leerse *de derecha a izquierda*, como la representación binaria de un número natural, el *número de la derecha* $\alpha_t$ (en cambio, si lo leemos de izquierda a derecha obtenemos una *secuencia infinita* de dígitos del sistema binario, que, después del último palote, son todos ceros). Llamaré $°\alpha_t$ ($°\rho_t$) a la representación binaria de $\alpha_t$ ($\rho_t$). Obsérvese que $\alpha_t$ es par si y sólo si $\square^t$ está en blanco y que $\rho_t$ es par si y sólo si está en blanco el cuadrado contiguo a la izquierda de $\square^t$. El tránsito de la etapa $t$ a la etapa $t$+1 puede describirse, como en nuestros diagramas, por un par de caracteres separados por dos puntos, $S$:$C$, donde $S \in \{1,0\}$ y $C \in \{1,0,\text{A},\text{R}\}$. Hay, pues, ocho casos posibles. Si el tránsito es 0:0 o 1:1, $\rho_{t+1} = \rho_t$ y $\alpha_{t+1} = \alpha_t$. En los seis casos restantes, no es difícil determinar los valores de $\rho_{t+1}$ y $\alpha_{t+1}$ en función de $\rho_t$, y $\alpha_t$, respectivamente. En el caso 0:1, $\rho_{t+1} = \rho_t$ y $\alpha_{t+1} = \alpha_t + 1$. En el caso 1:0, $\rho_{t+1} = \rho_t$ y $\alpha_{t+1} = \alpha_t - 1$. En el caso 0:A, $\rho_{t+1} = 2\rho_t$ ($°\rho_{t+1}$ es igual $°\rho_t$ seguido de un 0) y $\alpha_{t+1} = \alpha_t/2$ ($°\alpha_t$ es igual a $°\alpha_{t+1}$ seguido de un 0). En el caso 1:A, $\rho_{t+1} = 2\rho_t + 1$ ($°\rho_{t+1}$ es igual $°\rho_t$ seguido de un 1) y $\alpha_{t+1} = (\alpha_t-1)/2$ ($°\alpha_t$ es igual a $°\alpha_{t+1}$ seguido de un 1). En los casos 0:R y 1:R el resultado varía según que $\rho_t$ sea par o impar; en cambio, la paridad de $\alpha_t$ —dependiente de que $\square^t$ contenga un 0 o un 1— da lo mismo. Si $\rho_t$ es par, $\rho_{t+1} = \rho_t/2$ y $\alpha_{t+1} = 2\alpha_t$. Si $\rho_t$ es impar, $\rho_{t+1} = (\rho_t-1)/2$ y $\alpha_{t+1} = 2\alpha_t + 1$. Recordemos finalmente que en el sistema binario, una secuencia de $p$ unos representa el número $2^p - 1$, una secuencia de $p$ unos seguida de $q$ ceros representa a $(2^p - 1)2^q$, y una secuencia de $p$ unos seguida de un cero seguido de $q$ unos representa a $(2^p - 1)2^{q+1} + (2^q - 1)$. Por lo tanto, si $f$ es binaria ($r = 2$), cuando $\mathcal{M}_f$ computa a $f(x,y)$, y se aplica en la primera etapa de la computación al *par formado por $x$+1 palotes a la izquierda de un blanco a la izquierda de $y$+1 palotes (con el resto de la cinta en blanco), es claro que $\rho_0 = 0$ y $\alpha_0 = (2^{y+1} \dot{-} 1)2^{x+2} + (2^{x+1} \dot{-} 1)$. Adviértase que la función $\zeta_2:\langle x,y\rangle \mapsto (2^{y+1} \dot{-} 1)2^{x+2} + (2^{x+1} \dot{-} 1)$ es pr-recursiva.[44] En el caso general ($r \geq 1$),

44 La función pr-recursiva $\langle x,y\rangle \mapsto x \dot{-} y$ se definió en la p. 333, nota 22; es igual a $x - y$ si $y \leq x$; de otro modo, es igual a 0.

$$\alpha_0 = (2^{x_1+1} \dot{-} 1) + \sum_{j=1}^{r} (2^{x_j+1} \dot{-} 1)(2^{\Sigma_{k<j} x_k+2}) = \zeta_r(x_1, \ldots, x_r)$$

En la etapa $n$-ésima y final de la computación, $\mathcal{M}_f$ se aplica a una fila de $f(x)+1$ (en una cinta por lo demás en blanco), de modo que $\rho_n = 0$ y $\alpha_n = 2^{f(x)+1}$. Recordando lo dicho al final del párrafo 2, comprobamos que $\theta(\alpha_n) = f(x)$.

5. *Codificación de programas*. Sigamos considerando la MT $\mathcal{M}_f$ que computa la función $r$-aria $f$. Representamos numéricamente los estados, símbolos y operaciones de $\mathcal{M}_f$ de acuerdo con las convenciones siguientes: el estado $q_k$ ($0 \le k \le m$) estará representado por el número $k$+1; los símbolos $S_0$ (blanco) y $S_1$ (palote) estarán representados por 0 y 1, respectivamente; las operaciones 0, 1, A y R estarán representadas por 0, 1, 2, y 3, respectivamente. Las dos funciones binarias $O$ y $Q$ que definiré en seguida están completamente determinadas por el programa —o el diagrama— de $\mathcal{M}_f$ y lo representan de modo inequívoco. Ponemos $O(x,y) = u$ y $Q(x,y) = w$ si $\mathcal{M}_f$ ejecuta la operación representada por $u$ y pasa al estado representado por $w$ cuando lee el signo representado por $y$ en el estado representado por $x$. Pero si $x$ e $y$ no representan, respectivamente, un estado y un símbolo de $\mathcal{M}_f$, o si el programa de $\mathcal{M}_f$ no asocia una operación y un nuevo estado al estado y al símbolo representados por el par $\langle x,y \rangle$, entonces $O(x,y) = y$ y $Q(x,y) = 0$. Obsérvese que, en virtud de esta estipulación, si $x$ representa el estado final de $\mathcal{M}_f$, $Q(0,y) = 0$. Como el programa de $\mathcal{M}_f$ consta de un número finito de cuádruplos, es claro que $O$ y $Q$ son funciones *pr-recursivas* definidas mediante una lista finita de condiciones.

6. *La función $B_f$*. Definiré una función pr-recursiva ($r$+1)-aria $B_f$ con la siguiente propiedad: si $t$ es una etapa en la computación de $f(x)$ por $\mathcal{M}_f$ y $q(t)$ es el número que representa el estado de $\mathcal{M}_f$ en $t$ —pero $q(t) = 0$ si $\mathcal{M}_f$ se ha detenido antes de $t$— entonces $B_f(t,x) = [\rho_t, q(t), \alpha_t]$. Según la convención del párrafo 5, $q(0) = 1$. Recordando que $\rho_0 = 0$ y $\alpha_0 = \zeta_r(x)$, definimos:

(i) $$B_f(0,x) = [0,1,\zeta_r(x)]$$

Antes de seguir con la definición de $B_f$, introduzco algunos símbolos y abreviaturas. Sea $x/y = x$ dividido por $y$, si $x$ es divisible por $y$ y $x/y = 0$ si $x$ no

es divisible por $y$; Sea $\varepsilon:\mathbb{N} \rightarrow \{0,1\}$ tal que $\varepsilon(x) = 0$ si $x$ es par, $e(x) = 1$ si $x$ es impar. (Ambas funciones son pr-recursivas). Como vamos a definir $B_f$ de tal modo que $B_f(t,x)$ —para un dado $r$-tuplo $x$— sea el gödel de un triple, adoptaré tres abreviaturas para designar los tres números codificados en dicho gödel, para un dado $t$; llamaré $\mathfrak{r}_t$ al primero, $\mathfrak{q}_t$ al segundo y $\mathfrak{a}_t$ al tercero.[45] Llamaré $Q_t$ al número $Q(\mathfrak{q}_t,\varepsilon(\mathfrak{a}_t))$.[46] Obsérvese que ($a$) si $\mathfrak{q}_t = q(t)$ y $\mathfrak{a}_t = \alpha_t$, $Q_t$ es el número que representa el estado alcanzado por $\mathcal{M}_f$ en el tránsito de la etapa $t$ a la $t$+1, y ($b$) si $\mathcal{M}_f$ se detiene en o antes de la etapa $t$, $Q_t = 0$. Completamos la definición de $B_f$ con la siguiente estipulación:

$$\text{(ii)} \qquad B_f(t+1,\boldsymbol{x}) = \begin{cases} [\mathfrak{r}_t,Q_t,\mathfrak{a}_t] & si\ \varepsilon(\mathfrak{a}_t)=0\ y\ O(\mathfrak{q}_t,\varepsilon(\mathfrak{a}_t))=0 \\ [\mathfrak{r}_t,Q_t,\mathfrak{a}_t] & si\ \varepsilon(\mathfrak{a}_t)=1\ y\ O(\mathfrak{q}_t,\varepsilon(\mathfrak{a}_t))=1 \\ [\mathfrak{r}_t,Q_t,\mathfrak{a}_t+1] & si\ \varepsilon(\mathfrak{a}_t)=0\ y\ O(\mathfrak{q}_t,\varepsilon(\mathfrak{a}_t))=1 \\ [\mathfrak{r}_t,Q_t,\mathfrak{a}_t\dot{-}1] & si\ \varepsilon(\mathfrak{a}_t)=1\ y\ O(\mathfrak{q}_t,\varepsilon(\mathfrak{a}_t))=0 \\ [2\mathfrak{r}_t,Q_t,\mathfrak{a}_t/2] & si\ \varepsilon(\mathfrak{a}_t)=0\ y\ O(\mathfrak{q}_t,\varepsilon(\mathfrak{a}_t))=2 \\ [2\mathfrak{r}_t+1,Q_t,(\mathfrak{a}_t\dot{-}1)/2] & si\ \varepsilon(\mathfrak{a}_t)=1\ y\ O(\mathfrak{q}_t,\varepsilon(\mathfrak{a}_t))=2 \\ [\mathfrak{r}_t/2,Q_t,2\mathfrak{a}_t] & si\ \varepsilon(\mathfrak{a}_t)=0\ y\ O(\mathfrak{q}_t,\varepsilon(\mathfrak{a}_t))=3 \\ [(\mathfrak{r}_t\dot{-}1)/2,Q_t,2\mathfrak{a}_t+1] & si\ \varepsilon(\mathfrak{a}_t)=1\ y\ O(\mathfrak{q}_t,\varepsilon(\mathfrak{a}_t))=3 \\ 0 & en\ cualquier\ otro\ caso \end{cases}$$

La función $B_f$ es pr-recursiva, puesto que la hemos definido mediante una lista finita de condiciones pr-recursivas (cf. p. 411).[47] Las alternativas pro-

45 Estrictamente hablando, con el vocabulario del párrafo 3, ponemos $\mathfrak{r}_t = \Pi_1(B_f(t,z))$, $\mathfrak{q}_t = \Pi_2(B_f(t,z))$, $\mathfrak{a}_t = \Pi_3(B_f(t,z))$. Conforme a esta definición, $\mathfrak{r}_t = \mathfrak{q}_t = \mathfrak{a}_t = 0$ si $B_f(t,z) = 0$.

46 Dicho sin abreviaturas: $Q_t = Q(\Pi_2(B_f(t,z)),\varepsilon(\Pi_3(B_f(t,z))))$.

47 Imitando a Boolos y Jeffrey (1980, p. 94), he tomado un atajo para facilitar la lectura e interpretación de la cláusula (ii). En rigor, la definición de $B_f$ debiera ajustarse al esquema de recursión R5, con $g$ y $h$ definidas mediante una lista finita de condiciones pr-recursivas. Ponemos $g(x) = [0,1,\zeta_r(x)]$. En (ii), reemplazamos $B_f(t+1,x)$ por $h(t,y,x)$ y redefinimos $\mathfrak{r}_t$, $\mathfrak{q}_t$, $\mathfrak{a}_t$ y $Q_t$ como sigue: $\mathfrak{r}_t = \Pi_1(y)$, $\mathfrak{q}_t = \Pi_2(y)$, $\mathfrak{a}_t = \Pi_3(y)$, $Q_t = Q(\Pi_2(y),\varepsilon(\Pi_3(y)))$. Entonces, la función $B_f$ dada por

$$B_f(0,\boldsymbol{x}) = g(\boldsymbol{x})$$
$$B_f(t+1,\boldsymbol{x}) = h(t,B_f(t,\boldsymbol{x}),\boldsymbol{x})$$

es idéntica a la definida en el texto.

puestas en las primeras ocho líneas a la derecha de la llave corresponden precisamente a los ocho casos posibles examinados en el párrafo 4. Comparando lo dicho allí con los valores asignados a $B_f(t+1,z)$ en esas ocho líneas, comprobamos que, para cada $t \geq 0$, $B_f(t+1,x) = [\rho_{t+1},q(t+1),\alpha_{t+1}]$, conforme a nuestra exigencia inicial. Si en la computación de $f(x)$, $\mathcal{M}_f$ llega a su estado final en la etapa $t$, la lista $B_f(0,x), \ldots, B_f(t,x)$ provee una descripción completa, etapa por etapa, de dicha computación, y $B_f(y,x) = 0$ para todo $y > t$. Obviamente, $\mathfrak{q}_y = \Pi_2(B_f(y,x)) = 0$ si y sólo si $y > t$, pues si $y \leq t$, $\mathfrak{q}_y$ es el número que representa el estado en que se encuentra $\mathcal{M}_f$ en la etapa $y$, el cual, según nuestras convenciones, es siempre positivo.

Con estos recursos podemos, finalmente, establecer el resultado. Acabamos de ver que $\mathcal{M}_f$ se detiene en la etapa $t$ de la computación de $f(x)$ si y sólo si $t = \mu y(\Pi_2(B_f(y+1,x)) = 0)$. La función $\phi{:}x \mapsto \mu y(\Pi_2(B_f(y+1,x)) = 0)$ es recursiva (aunque no necesariamente pr-recursiva). Como indiqué al final del párrafo 4, si $\mathcal{M}_f$ se detiene en la etapa $t$ de la computación de $f(x)$, $f(x) = \theta(\alpha_t)$ =. Pero $\alpha_t = \mathrm{a}_t = \Pi_3(B_f(t,x))$. Por lo tanto,

$$f(x) = \theta(\Pi_3(B_f(\phi(x),x)) \qquad (*)$$

Como función compuesta de funciones recursivas, $f$ es ciertamente recursiva.

*III. Forma normal de una función recursiva.* Sea $f$ una función recursiva $r$-aria. Entonces, $f$ es T-computable (I). Por lo tanto, $f$ es idéntica a la función compuesta al lado derecho de la ecuación (*). Dicha función se forma por composición de la función pr-recursiva $\theta \circ \Pi_3 \circ B_f$ con la función $\phi$ definida por minimización (según el esquema R6) a partir de una función pr-recursiva. Por lo tanto, toda función recursiva puede definirse por una serie finita de aplicaciones de los esquemas de recursión primitiva R1-R5, con *a lo sumo una* aplicación del esquema R6.[48]

[48] El resultado III no es igual al Teorema de la Forma Normal de Kleene (1936) enunciado en la p. 366 (inmediatamente antes de la nota 5). Kleene mostró que cualquier función recursiva general según la definición de Herbrand-Gödel tiene una forma normal como la descrita arriba, y *basó en esto* la conclusión de que las funciones recursivas generales de Herbrand-Gödel pueden todas definirse mediante los esquemas R1-R6. Pero nuestro resultado I, invocado en la prueba del resultado III, se basa él mismo en la *definición* de función recursiva mediante los esquemas R1-R6.

*IV. La máquina universal de Turing*. Turing (1936) da el programa de una máquina que, aplicada al gödel de una MTR cualquiera, computa el número real computable por ésta. La tarea de esa máquina consiste, esencialmente, en descifrar el gödel a que se aplica y seguir el programa allí encapsulado. Es probable que este concepto de una máquina universal haya inspirado a von Neumann (hacia 1945) la idea matriz de la computadora moderna: en vez de materializar el programa en los circuitos del aparato (como *hardware*), suministrárselo como información (en la forma de *software*). A continuación esbozaré una prueba de que, para cada entero positivo $r$ hay una MT universal $\mathcal{U}_r$ que computa el valor de cualquier función T-computable $r$-aria en cada $r$-tuplo $x$ perteneciente a su dominio, cuando se le suministra el $(r+1)$-tuplo $\langle i,x\rangle$, donde $i$ es el gödel de la MT que computa a $f$.[49] Llamaré $\mathcal{M}_i$ a la MT cuyo gödel es $i$ y $f_i$ a la función que ella computa. Diré que $i$ es el *indice* de la función $f_i$. El índice $i$ de la función T-computable $f_i$ encapsula el programa $\mathcal{M}_i$ y, por ende, contiene toda la información necesaria para definir la función pr-recursiva $B_{f_i}$ descrita en II.6; en adelante, escribo $B_i$ en vez de $B_{f_i}$. Si $f_i$ es $r$-aria y el $r$-tuplo $x$ pertenece a su dominio, la computación de $f_i(x)$ por $\mathcal{M}_i$ termina en cierta etapa $t$ y —como se vio en II.6— la secuencia finita $B_i(0,x), \dots , B_i(t,x)$ contiene toda la información necesaria para describir dicha computación, etapa por etapa. Si conocemos $t$ esta información puede encapsularse en el gödel $[B_i(0,x), \dots , B_i(t,x)]$. Como $B_i(u,x) \neq 0$ para todo $u \leq t$ y $B_i(u,x) = 0$ para todo $u > t$, en verdad no hace falta conocer $t$ para definir un gödel que encapsule la información. Ponemos, simplemente,[50] $\Gamma(i,\boldsymbol{x}) = \prod_{k=0}^{\infty} (p_{k+1})^{B_i(k,\boldsymbol{x})}$. Si $f_i$ está definida en $x$, ocurre que a partir de cierto valor finito del índice $k$ los factores del producto infinito $\Gamma(i,x)$ llevan todos el exponente 0 y son, por ende, iguales a 1. En tal caso, $\Gamma(i,x)$ es precisamente el gödel de la computación de $f_i(x)$ por $\mathcal{M}_i$, tal como se lo definió arriba. Si $f_i$ no está definida en $x$, $\Gamma(i,x)$ excede a cualquier número señalable y puede considerarse indefinido. Recordando la definición 7 en el paso 4º de la Sección 2.10.2, vemos que, si $\Gamma(i,x)$ está definido, $\ell(\Gamma(i,x))$ es igual al número de los factores primos de $\Gamma(i,x)$ y, por ende, es igual al número $t$ de

49 El argumento bosquejado a continuación se debe a Kleene; *vide* p. 410, nota 43.

50 Recuérdese que, según la definición 5 en la p. 336, $p_0 = 0$ y, para $r > 0$, $p_r$ es el $r$-ésimo primo ($p_1 = 2$, $p_2 = 3$, $p_3 = 5, \dots$ ).

la etapa en que termina la computación de $f_i(x)$ por $\mathcal{M}_i$. Definimos la función ternaria $T_r$ mediante una lista de dos condiciones:

| | |
|---|---|
| $T_r(u,v,w) = 0$ | si $u$ es el índice de una función T-computable $r$-aria, $v$ es el gödel de un $r$-tuplo $v$ en el dominio de $f_u$ y $w = \ell(\Gamma(u,v))$. |
| $T_r(u,v,w) = 1$ | en cualquier otro caso. |

Como las condiciones son recursivas, $T_r$ es una función recursiva.[51] También es recursiva la función binaria $\psi:\langle u,v\rangle \mapsto \mu w(T_r(u,v,w) = 0)$, cuyo valor, en cada par $\langle u,v\rangle$ en que esté definida, es precisamente el número de la etapa en que la MT $\mathcal{M}_u$ se detiene en la computación de $f_u(v)$. Recordando la ecuación (*) en la p. 415, defino la función $(r+1)$-aria

$$\Phi_r:\langle i,x\rangle \mapsto \theta(\Pi_3(B_i(\psi(i,[x]),x)).$$

$\Phi_r$ es una función recursiva y, por ende, T-computable. Sea $\mathcal{U}_r$ la MT que computa a $\Phi_r$. $\mathcal{U}_r$ es la MT buscada.

*V. El problema de la detención es insoluble*. El problema de la detención puede formularse así: decidir mediante un algoritmo si una MT se detiene cuando se aplica a cierto *número. Si aceptamos la tesis de Church (o el análisis de la actividad de un calculista propuesto por Turing), el problema equivale a este otro: hallar una función T-computable binaria $\Delta$ tal que $\Delta(u,x) = 0$ si $u$ es el índice de una función T-computable $f_u$ definida en $x$, y $\Delta(u,x) = 1$ en cualquier otro caso. Evidentemente $\Delta(u,x) = 0$ si y sólo si la computación de $f_u(x)$ termina en alguna etapa $t$, esto es, si y sólo si $\exists w(T_1(u,[x],w) = 0)$. Mostraré que el problema de la detención es insoluble o, mejor dicho, que tiene una solución negativa: la función T-computable $\Delta$ *no existe*. Con ese propósito, defino la siguiente función total:

[51] La recursividad de las condiciones impuestas a $u$ y $v$ puede establecerse utilizando los predicados y funciones definidos en el paso 4º de nuestra demostración del primer teorema de incompletud de Gödel (pp. 335-341). Pero también podemos invocar la Tesis de Church, puesto que hay sin duda algoritmos para decidir si un número dado es o no el índice de una función T-computable o el gödel de un $r$-tuplo (para un $r > 0$ fijo). La condición impuesta a $w$ depende de las condiciones anteriores y de las funciones recursivas $\ell$ y $\Gamma$.

$$\varphi(x) = f_x(x) + 1 \qquad \text{si } \exists w(T_1(x,[x],w) = 0)$$
$$\varphi(x) = 0 \qquad \text{en cualquier otro caso.}$$

Si φ es T-computable, tendrá un índice $a$. En otras palabras, φ será idéntica a la función T-computable unaria $f_a$. Como φ es total, está definida en $a$, de modo que $\exists w(T_1(a,[a],w) = 0)$. Por lo tanto, $f_a(a) = \varphi(a) = f_a(a) + 1$. Como esto es imposible, concluimos que φ no es T-computable. Supongamos ahora que existe la función T-computable Δ arriba descrita. Entonces, la función φ podría computarse así: dado un número $x$, computar $\Delta(x,x)$; si $\Delta(x,x) = 0$, $x$ es el índice de la función T-computable $f_x$ definida en $x$ y, por ende, podemos computar $f_x(x) + 1$ que es el valor correcto de $\varphi(x)$ en ese caso; si $\Delta(x,x) = 1$, ponemos $\varphi(x) = 0$, que es su valor correcto en este caso. Como φ no puede ser T-computable, so pena de contradicción, tampoco puede existir una función T-computable Δ como la arriba descrita.

*VI. El cálculo predicativo de primer orden es indecidible*. La solución negativa del problema de la detención es homóloga, en la teoría de las MT, al Teorema XVIII de Church 1936 (Sección 2.11.2), en la teoría de la λ-computabilidad. Por eso, no ha de sorprendernos que ella lleve directamente a la solución negativa del *Entscheidungsproblem* de Hilbert y Ackermann. Por su parte, Turing (1936), como era natural, derivó este resultado de su solución negativa del problema de la circularidad, concerniente a sus MTRs. El *Entscheidungsproblem* de Hilbert y Ackermann demanda un algoritmo para determinar si una fórmula dada del cálculo predicativo de primer orden es válida o no. En virtud del Teorema de Completud de Gödel (1929, 1930) estudiado en el Capítulo 2.8, el problema tiene una solución positiva si y sólo si hay un algoritmo para determinar si dicha fórmula es deducible o no en el cálculo organizado como sistema deductivo completo. Por lo tanto, según la Tesis de Church, el *Entscheidungsproblem* sólo podrá resolverse positivamente si existe una función computable δ tal que (i) $\delta(x)$ está definida si y sólo $x$ es el gödel de una fórmula ξ del cálculo predicativo de primer orden y (ii) $\delta(x) = 0$ si ξ es deducible en el cálculo y $\delta(x) = 1$ si ξ no es deducible en el cálculo. Este problema está íntimamente ligado al problema de la detención porque la teoría de las funciones T-computables se puede formalizar en el cálculo predicativo de primer orden. En particular, es posible asociar a cada par de números $\langle x,y\rangle$, una fórmula $\xi(x,y)$ que es deducible en el cálculo

si y sólo si $x$ es el índice de una función T-computable unaria definida en $y$. La función que asigna a $\langle x,y \rangle$ el gödel de la fórmula correspondiente $\xi(x,y)$ es computable. Llamémosla $\gamma$. Es claro que, si existiera la función computable $\delta$ arriba descrita tendríamos que $\delta(\gamma(x,y)) = 0$ si y sólo si $\exists w(T_1(x,[y],w) = 0)$. Por consiguiente, la función compuesta $\delta \circ \gamma$ sería idéntica a la función $\Delta$ que resuelve el problema de la detención (véase el V). Pero hemos demostrado que esa función no existe. Por lo tanto, $\delta$ no existe.[52]

[52] Varios autores dan la demostración en detalle. Especialmente clara y simple es la versión de Boolos y Jeffrey (1980, Capítulo 10), basada en Büchi 1962.

## 2.12 CONSISTENCIA DE LA ARITMÉTICA: LA PRUEBA DE GENTZEN

Los escritos de Gödel, Church, Kleene y Turing a que me refiero en los Capítulos 2.10 y 2.11 son obras maestras de la "metamatemática", que aplican con ejemplar virtuosismo los recursos de la aritmética finitista (Capítulo 2.5) al estudio de cálculos lógicos. Sus resultados no favorecen las aspiraciones de la escuela de Hilbert. Hilbert y Bernays (1939) anuncian a sus lectores que, a la luz del descubrimiento de Gödel, su programa de fundamentación de las matemáticas demanda una ampliación del punto de vista finitista.[1] Toda la ampliación requerida consiste en admitir como un método propiamente "finito" la inducción transfinita que Gerhard Gentzen invoca en sus dos demostraciones de la consistencia de la aritmética formalizada (1935, 1938).

Las dos demostraciones de Gentzen difieren significativamente en los detalles —el cálculo a que se aplica la segunda difiere del considerado en la primera— pero ambas se inspiran en la misma idea. Se propone un sistema deductivo para lo que Gentzen llama "teoría pura de los números" (*reine Zahlentheorie*), que es lo mismo que aquí hemos estado llamando aritmética elemental. Imitando a Gentzen, llamaré *derivaciones* a las pruebas de ese sistema (él dice *Herleitungen*). Las derivaciones admisibles se ordenan por niveles, a partir de un nivel mínimo. El orden está diseñado de tal modo que, si hay una derivación $\mathbf{\Delta}$ que concluye en una contradicción, tiene que haber otra derivación $\mathbf{\Delta}^*$ con la misma conclusión, tal que $\mathbf{\Delta}^*$ tiene un nivel más bajo que $\mathbf{\Delta}$. Por otra parte, será obvio que las derivaciones de nivel mínimo no pueden concluir en una contradicción. De esto se deduce, por

1 Bernays reconoce, en el prefacio al tomo II de *Fundamentos de la matemática*: "En contraste con la definición anterior del 'punto de vista finitista', ha resultado necesario ampliar el marco de los modos de inferencia sustantivos admitidos en la teoría de la prueba" ("die Notwendigkeit [hat sich] ergeben [ . . . ], den Rahmen der für die Beweistheorie zugelassenen inhaltlichen Schlußweisen gegenüber der vorherigen Abgrenzung des 'finiten Standpunktes' zu erweitern"—Hilbert y Bernays 1939, p. vii).

inducción sobre el nivel de las derivaciones, que ninguna derivación del sistema puede desembocar en una contradicción. Pero la inducción de que aquí se trata rebasa los límites del "modo recursivo de pensar" patrocinado por Skolem (1923) y aceptado sin reservas en la escuela de Hilbert. Aunque el conjunto de las derivaciones admisibles es, por cierto, numerable,[2] el orden que hay que darle para los efectos del argumento de Gentzen no es isomórfico a $\omega$, sino a un ordinal mucho mayor.[3] Por lo tanto, la inducción requerida se extiende sobre todos los ordinales menores que ése. Se razona así: (i) Ninguna derivación del nivel ínfimo lleva a una contradicción. (ii) Si una derivación de nivel $\zeta$ lleva a una contradicción, hay una derivación de un nivel $\eta < \zeta$ que también lleva a una contradicción. (iii) Por lo tanto, si ninguna derivación de nivel $\eta < \zeta$ lleva a una contradicción, las derivaciones de nivel $\zeta$ tampoco llevan a una contradicción. Esta forma de razonar no me merece ninguna duda, pero me cuesta trabajo llamarla *finita* o *finitista*.

La argumentación de Gentzen es tortuosa —tanto, que juzgó necesario reescribirla— pero, exceptuando la inducción descrita, es enteramente elemental; en particular, la construcción, para cada derivación con conclusión contradictoria, de otra derivación equivalente de menor nivel es una simple transformación de un objeto finito en otro objeto finito, y la complicación viene sólo de la variedad de los casos posibles. Examinaré en detalle la versión de 1938. Presento el cálculo en la Sección 2.12.1. Este cálculo es inconsistente si y sólo si hay derivaciones de la clase que llamaré "fatales". En la Sección 2.12.2 demuestro que toda derivación fatal se deja reducir a otra de la misma clase, construida de tal modo que preceda a la primera en el orden definido por Gentzen. En la Sección 2.12.3 doy la definición de este orden, lo uso para fundamentar la inducción arriba esbozada y hago algunas observaciones sobre la índole de ésta y su utilidad dentro y fuera del programa de Hilbert.

2 Como cada derivación es una colección finita de signos es posible asignarle un gödel a cada una, ordenarlos de menor a mayor y contarlos.

3 Concretamente, el primer ordinal $\xi$, tal que $\omega^{\xi} = \xi$. Este es el número que Cantor llamó $\varepsilon_0$ (1895/97, §20). Es, por cierto, un número de la Clase II: el conjunto de sus predecesores es infinito pero numerable (véase el capítulo 1.5).

### *2.12.1 Un cálculo aritmético*

En 1934, Gentzen y Jaśkowski introdujeron en sendas publicaciones independientes los sistemas de "deducción natural" que prevalecen hoy en la enseñanza de la lógica. En el Apéndice IX.G propuse, a modo de ejemplo, un sistema deductivo de ese tipo, adaptado a nuestro CP1. El lector que no lo haya visto debiera echarle una mirada ahora (pp. 496-99). Ese sistema es muy parecido al que Gentzen adoptó en su artículo de 1935. En 1938 adopta otro, que presentaré aquí. El cambio facilita la reducción en el nivel de las derivaciones. Pero el nuevo cálculo es por lo menos tan poderoso como el anterior. En su parte lógica, es por lo menos tan poderoso como el cálculo de primer orden que Gödel (1930) demostró que era completo (véase Apéndice XVIII). Como se trata de probar la consistencia de la aritmética formalizada en el cálculo, no importa que sea *más* poderoso: si no se puede derivar una contradicción en el cálculo más poderoso, tampoco se la puede derivar en el más débil.

El cálculo de Gentzen (1938) es un cálculo predicativo de primer orden. Gentzen no enuncia reglas para la formación de variables y predicados, pero podemos adoptar las de nuestro CP1 (Apéndice IX.A). El signo 1 es la única constante (en la aritmética formalizada, '1' designa el primer número natural).[4] Cada variable es un *término*. 1 es un *término*. Si $\tau$ es un término, $\tau'$ es un *término*. (El signo $'$ opera pues como un functor pospuesto a su argumento; en la aritmética formalizada, $\tau'$ designa el siguiente del número designado por $\tau$). Un término que no contiene una variable es un *término numérico*. Se admite cualquier número de predicados de cualquier *n*-ariedad con la única condición de que sean decidibles: si $\Phi$ es un predicado *n*-ario y $\tau_1, \ldots, \tau_n$ son términos en que no figura ninguna variable, tiene que haber un procedimiento efectivo para decidir si $\Phi\tau_1 \ldots \tau_n$ es o no verdadero (en la interpretación aritmética indicada). En los ejemplos de Gentzen figuran los

[4] La elección del signo indica, sin duda, que Gentzen prefería contar desde uno: *eins, zwei, drei* . . . Dadas nuestras preferencias, habría que entender que '1' designa el número cero (como en las representaciones gráficas de máquinas de Turing en la Sección 2.11.4), o reemplazarlo por el signo '0'. Pero el lector ya se habrá acostumbrado a pensar que todo esto da lo mismo; el significado de la constante '1' queda fijado tan precisamente como es posible si estipulamos que ésta designa *el único número que no es el siguiente de otro*.

predicados binarios = y >, escritos, como es habitual, entre los términos a que se aplican.[5] El procedimiento de decisión en ambos casos es obvio: si $\tau_1$ y $\tau_2$ son términos sin variables, $\tau_1 = \tau_2$ es verdadero sí y sólo si hay igual número de signos ′ en $\tau_1$ y $\tau_2$ y $\tau_1 > \tau_2$ es verdadero sí y sólo si hay mayor número de signos ′ en $\tau_1$ que en $\tau_2$.[6] Conviene advertir que '=' *no es* aquí un signo *lógico*: no hay reglas de inferencia para introducirlo ni su presencia o ausencia desempeña un papel en las reducciones de nivel. Llamaré como siempre *fórmula elemental* a la que consta de un predicado y el número justo de términos (Gentzen dice *Primformel*).

Igual que en el cálculo del Apéndice IX.G, las fórmulas del cálculo se organizan en secuentes. Un secuente está formado por dos listas finitas (posiblemente vacías) de fórmulas separadas por comas; entre las dos listas va el signo ➛. Tal como en el Apéndice IX.G, llamo prefórmulas a las fórmulas que preceden a ➛ y posfórmulas a las que le siguen. A diferencia del cálculo estudiado allí, cada secuente puede contener más de una posfórmula. Un secuente sin variables libres es inválido (i) si todas sus prefórmulas son verdaderas y todas sus posfórmulas son falsas en cualquier interpretación en que la constante 1 designa el primer número natural y el término $\tau'$ designa el siguiente del número designado por $\tau$; (ii) si no tiene prefórmulas ni posfórmulas, esto es, si se trata del secuente "vacío". En cualquier otro caso, el secuente es válido. En otras palabras: la *conjunción* de las prefórmulas de un secuente válido garantiza la *disyunción* de las posfórmulas.

Hay seis símbolos lógicos. Como acabo de indicar, el signo ➛ separa las prefórmulas de las posfórmulas en un secuente. Los otros cinco —que, imitando a Gentzen, llamaré *conectivos* (él dice *Verknüpfungszeichen*)— son la negación ¬, la conjunción ∧,[7] la disyunción ∨ y los signos de cuantificación ∀ y ∃. Dichos conectivos tienen su significado habitual.[8] Las fórmulas no

[5] Para ajustarnos estrictamente al CP1 tendríamos que elegir dos predicados binarios, digamos $P^2$ y $P^2{}_1$ y estipular que $(\tau_1 = \tau_2)$ abrevia a $P^2\tau_1\tau_2$ y $(\tau_1 > \tau_2)$ abrevia a $P^2{}_1\tau_1\tau_2$.

[6] Fuera de = y >, veo un predicado más, el predicado ternario '*x* es congruente con *y* módulo *z*', que se utiliza una sola vez, en la fórmula '$1''' \equiv 1 \pmod{1''}$', presentada como ejemplo ilustrativo (Gentzen 1938, p. 23, al final del §1.4).

[7] En vez de ∧, Gentzen usa & como signo de conjunción.

[8] Conviene, sí, tener presente la siguiente advertencia de Gentzen (1935, p. 526): si $\forall\mathfrak{x}\mathfrak{F}(\mathfrak{x})$ es una fórmula con un solo cuantificador y sin otra variable que $\mathfrak{x}$, "no tenemos que asociar a ese ∀ la representación de una cantidad infinita *cerrada* de aseveraciones

elementales se construyen a partir de las elementales, mediante conectivos y paréntesis, tal como en el CP1.

Para construir los secuentes y las derivaciones se requieren tres nuevos signos de puntuación: la coma, que separa dos prefórmulas o dos posfórmulas de un secuente; el espacio, que separa dos premisas de una inferencia, y la raya horizontal, que va en cada inferencia entre la o las premisas y la conclusión. El espacio que es un signo de puntuación del cálculo se distingue de otros blancos en la página porque deslinda a derecha e izquierda con secuentes y abajo con una raya.[9] (A diferencia de lo que ocurre en nuestro Apéndice IX.G, el espacio y la raya no son aquí signos auxiliares introducidos para facilitar nuestra presentación *en castellano* de las reglas de inferencia, sino que forman parte de las derivaciones y contribuyen a determinar su estructura; son, pues, *signos del cálculo*).

Una *inferencia* (Gentzen dice *Schlußfigur*, "figura de inferencia") consta de uno o dos secuentes escritos sobre una raya horizontal y un secuente escrito bajo esa raya. Aquéllos son las *premisas*; éste la *conclusión* (Gentzen los llama *Obersequenzen* —"secuentes de arriba"— y *Untersequenz* —"secuente de abajo"—, respectivamente). Cada inferencia se ajusta a uno de los veinte esquemas presentados a continuación. Siguiendo a Gentzen, represento mediante una mayúscula griega una fila —posiblemente vacía— de fórmulas separadas por comas; mediante una mayúscula gótica, una fórmula cualquiera; la minúscula gotica $\mathfrak{t}$ representa un término cualquiera, las minúsculas góticas $\mathfrak{a}$ y $\mathfrak{x}$ representan variables, y los signos lógicos y de puntuación están representados homónimamente por ellos mismos. $\mathfrak{F}(\mathfrak{t}/\mathfrak{x})$ es la fórmula que se obtiene al reemplazar la variable $\mathfrak{x}$ por el término $\mathfrak{t}$ en todas las posiciones donde $\mathfrak{x}$ está libre en $\mathfrak{F}$ (cf. Apéndice IX.A, p. 485). La variable representada en algunos esquemas por la letra $\mathfrak{a}$ está sujeta a las restricciones que se indican; Gentzen la llama la *variable propia* (*Eigenvariable*) de la inferencia respectiva. En la inferencia estructural llamada *corte*, las fórmulas

particulares, sino que podemos concebir su sentido '*finitamente*' ('*finit*') así: 'Si la $\mathfrak{x}$ se reemplaza sucesivamente por números, empezando con el 1, entonces, por mucho que se avance en la formación de números, se obtiene en cada caso una aseveración verdadera'."

9 Si esta caracterización no parece suficientemente precisa, se puede reemplazar el espacio entre premisas por un signo impreso, por ejemplo, el punto y coma.

iguales, representadas con $\mathfrak{D}$, son las *fórmulas cortadas* en el corte. En las inferencias con conectivo, la fórmula con conectivo destacada en la conclusión es la *fórmula principal* de la inferencia.

ESQUEMAS DE LAS INFERENCIAS ESTRUCTURALES[10]

Debilitamiento

$$\frac{\Gamma \rightarrow \Theta}{\mathfrak{D}, \Gamma \rightarrow \Theta} \qquad \frac{\Gamma \rightarrow \Theta}{\Gamma \rightarrow \Theta, \mathfrak{D}}$$

Contracción

$$\frac{\mathfrak{D}, \mathfrak{D}, \Gamma \rightarrow \Theta}{\mathfrak{D}, \Gamma \rightarrow \Theta} \qquad \frac{\Gamma \rightarrow \Theta, \mathfrak{D}, \mathfrak{D}}{\Gamma \rightarrow \Theta, \mathfrak{D}}$$

Permutación

$$\frac{\Delta, \mathfrak{D}, \mathfrak{E}, \Gamma \rightarrow \Theta}{\Delta, \mathfrak{E}, \mathfrak{D}, \Gamma \rightarrow \Theta} \qquad \frac{\Gamma \rightarrow \Theta, \mathfrak{D}, \mathfrak{E}, \Delta}{\Gamma \rightarrow \Theta, \mathfrak{E}, \mathfrak{D}, \Delta}$$

Corte

$$\frac{\Gamma \rightarrow \Theta, \mathfrak{D} \qquad \mathfrak{D}, \Delta \rightarrow \Lambda}{\Gamma, \Delta \rightarrow \Theta, \Lambda}$$

[10] A la inferencia estructural que llamo debilitamiento, Gentzen la llama *Verdünnung*, esto es, 'dilución' o 'desleímiento' ('thinning' en la traducción inglesa de Szabo); pero, como luego señala, "un 'desleímiento' representa tan sólo un *debilitamiento* (*Schwächung*) del sentido sustantivo de un secuente" (1938, p. 31; cursiva en el original).

Esquemas de las inferencias con conectivos

| | |
|---|---|
| $\wedge$ | $\vee$ |
| $\dfrac{\Gamma \to \Theta, \mathfrak{A} \qquad \Gamma \to \Theta, \mathfrak{B}}{\Gamma \to \Theta, \mathfrak{A} \wedge \mathfrak{B}}$ | $\dfrac{\mathfrak{A}, \Gamma \to \Theta \qquad \mathfrak{B}, \Gamma \to \Theta}{\mathfrak{A} \vee \mathfrak{B}, \Gamma \to \Theta}$ |
| $\dfrac{\mathfrak{A}, \Gamma \to \Theta}{\mathfrak{A} \wedge \mathfrak{B}, \Gamma \to \Theta} \qquad \dfrac{\mathfrak{B}, \Gamma \to \Theta}{\mathfrak{A} \wedge \mathfrak{B}, \Gamma \to \Theta}$ | $\dfrac{\Gamma \to \Theta, \mathfrak{A}}{\Gamma \to \Theta, \mathfrak{A} \vee \mathfrak{B}} \qquad \dfrac{\Gamma \to \Theta, \mathfrak{B}}{\Gamma \to \Theta, \mathfrak{A} \vee \mathfrak{B}}$ |
| $\forall$ | $\exists$ |
| $\dfrac{\Gamma \to \Theta, \mathfrak{F}(\mathfrak{a}/\mathfrak{x})}{\Gamma \to \Theta, \forall \mathfrak{x}\mathfrak{F}}$ | $\dfrac{\mathfrak{F}(\mathfrak{a}/\mathfrak{x}), \Gamma \to \Theta}{\exists \mathfrak{x}\mathfrak{F}, \Gamma \to \Theta}$ |
| si la variable $\mathfrak{a}$ no figura en la conclusión | si la variable $\mathfrak{a}$ no figura en la conclusión |
| $\dfrac{\mathfrak{F}(\mathfrak{t}/\mathfrak{x}), \Gamma \to \Theta}{\forall \mathfrak{x}\mathfrak{F}, \Gamma \to \Theta}$ | $\dfrac{\Gamma \to \Theta, \mathfrak{F}(\mathfrak{t}/\mathfrak{x})}{\Gamma \to \Theta, \exists \mathfrak{x}\mathfrak{F}}$ |
| $\neg$ | |
| $\dfrac{\mathfrak{A}, \Gamma \to \Theta}{\Gamma \to \Theta, \neg\mathfrak{A}}$ | $\dfrac{\Gamma \to \Theta, \mathfrak{A}}{\neg\mathfrak{A}, \Gamma \to \Theta}$ |

Esquema de las inferencias inductivas

IM

$$\frac{\mathfrak{F}(\mathfrak{a}/\mathfrak{x}),\ \Gamma \rightarrowtail \Theta,\ \mathfrak{F}(\mathfrak{a}'/\mathfrak{x})}{\mathfrak{F}(1/\mathfrak{x}),\ \Gamma \rightarrowtail \Theta,\ \mathfrak{F}(\mathfrak{t}/\mathfrak{x})}$$

si la variable $\mathfrak{a}$ no figura en la conclusión

Como puede verse, éste último esquema autoriza la inducción matemática ordinaria como procedimiento normal de inferencia dentro del cálculo. El lector debe cerciorarse de que los otros 19 esquemas llevan en todo caso de premisas válidas a conclusiones válidas.

Una derivación es una colección finita secuentes, separados por rayas horizontales o espacios, con las propiedades que se enumeran a continuación.

(i) Cada raya horizontal separa las premisas de la conclusión de una inferencia ajustada a uno de los 20 esquemas.
(ii) Hay un y sólo un *secuente final* que no está sobre ninguna raya (no es premisa de una inferencia).
(iii) Hay uno o más *secuentes iniciales* que no están bajo ninguna raya (no son conclusiones de inferencias).
(iv) Los secuentes iniciales pertenecen a una de las dos clases de *secuentes básicos*, definidas a continuación:
  (*a*) si $\mathfrak{F}$ es una fórmula cualquiera, $\mathfrak{F} \rightarrowtail \mathfrak{F}$ es un *secuente básico lógico* (un SBL);
  (*b*) si $\Gamma$ y $\Delta$ son listas finitas (no vacías) de fórmulas elementales, el secuente $\Gamma \rightarrowtail \Delta$ es un *secuente básico matemático* (un SBM) si *todo* secuente obtenido reemplazando uniformemente cada variable en $\Gamma$ y $\Delta$ por un término numérico es un secuente válido.

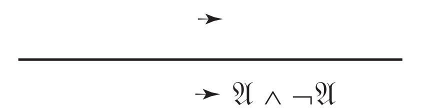
Según esto, una derivación puede considerarse como un árbol finito de secuentes, ordenado parcialmente por la relación $x < z$ definida recursivamente como sigue: $x < z$ (i) si $x$ es una premisa y $z$ es la conclusión de una misma inferencia o (ii) si hay en la derivación un secuente $u$ tal que $x < u$ y $u < z$.

El Apéndice XVIII contiene varios ejemplos de derivaciones que el lector puede examinar a la luz de lo que llevamos dicho.

En la caracterización de las derivaciones el único ingrediente que requiere una explicación es la definición de los SBMs. Para apreciar su alcance, conviene distinguir entre SBMs con y sin variables. Como ya indiqué, todos los predicados del cálculo son decidibles. Por lo tanto, si $\Gamma$ y $\Delta$ son listas finitas de fórmulas elementales sin variables, habrá un algoritmo para decidir si el secuente $\Gamma$ ➛ $\Delta$ es válido o no, esto es, si $\Gamma$ ➛ $\Delta$ es o no un SBM. Pero si hay una o más variables en $\Gamma$ o en $\Delta$, tal decisión es generalmente impracticable, ya que las sustituciones posibles de esa o esas variables por términos numéricos no pueden completarse en un número finito de pasos y la validez del secuente no puede certificarse sin ensayarlas todas. Como los SBMs operan, en rigor, como axiomas del cálculo, *éste no es lo que llamamos un cálculo efectivo*. Ello no obstaculiza la tarea que Gentzen se ha propuesto. Al contrario: mientras más poderoso sea un cálculo, mayor es el riesgo de que sea inconsistente. Por eso, la generosa aceptación de toda clase de axiomas aritméticos, decidibles o no, realza el valor de la prueba de consistencia.

### *2.12.2 Reducciones*

El sistema deductivo descrito es inconsistente si y sólo si permite derivar el secuente vacío ‘ ➛ ’, que no tiene prefórmulas ni posfórmulas. En efecto, si ‘ ➛ ’ es el secuente final de una derivación y $\mathfrak{A}$ es una fórmula cualquiera, hay una derivación cuyo secuente final es ‘ ➛ $\mathfrak{A} \wedge \neg\mathfrak{A}$’; esta derivación se obtiene simplemente añadiendo una raya horizontal y el secuente ‘ ➛ $\mathfrak{A} \wedge \neg\mathfrak{A}$’ bajo la derivación anterior. (La inferencia añadida



es una inferencia estructural por debilitamiento). Por otra parte, si ‘ ➛ $\mathfrak{A}$’ y ‘ ➛ $\neg\mathfrak{A}$’ son los secuentes finales de dos derivaciones del cálculo, se las

puede unir en una derivación cuyo secuente final es ‘ $\rightarrowtail$ ’:

$$\dfrac{\rightarrowtail \neg\mathfrak{A} \qquad \dfrac{\rightarrowtail \mathfrak{A}}{\neg\mathfrak{A} \rightarrowtail}\ (\neg)}{\rightarrowtail}\ (\text{corte})$$

Llamaré, por eso, *fatal* (entiéndase: para la consistencia del cálculo) a cualquier derivación cuyo secuente final es ‘ $\rightarrowtail$ ’.[11] Probaremos que no puede haber una derivación fatal. En la prueba se usarán algunos términos que defino a continuación.

Si una fórmula empieza con un conectivo, éste es el *conectivo principal* de esa fórmula; si empieza con un paréntesis, su *conectivo principal* es el conectivo cuyo alcance incluye el signo que sigue inmediatamente a este paréntesis.[12]

El *grado* de una fórmula es el número de conectivos que contiene. El grado de un corte es el grado de las fórmulas cortadas en él (representadas por $\mathfrak{D}$ en el esquema). El grado de una inferencia inductiva es el grado de la fórmula a que se refiere la inducción (representada por $\mathfrak{F}$ en el esquema).

Sea $\boldsymbol{\Delta}$ una derivación cualquiera. Decimos que el secuente σ *precede inmediatamente* en $\boldsymbol{\Delta}$ al secuente σ′ (y que σ′ *sigue inmediatamente* a σ) si σ es una premisa de una inferencia en $\boldsymbol{\Delta}$ cuya conclusión es σ′. Si $\sigma < \sigma'$ (en el sentido definido en 2.12.1), decimos que σ precede a σ′ o que está sobre σ′ y que σ′ sigue a σ o está bajo σ.

Asignaremos una *altura* a cada secuente σ en la derivación $\boldsymbol{\Delta}$. Para ello, atendemos al grado de cada corte e inferencia inductiva cuya conclusión está bajo σ en $\boldsymbol{\Delta}$. El mayor de esos grados es la *altura* de σ en $\boldsymbol{\Delta}$. Este número se utiliza luego de un modo decisivo al ordenar las derivaciones.

11 Gentzen la llama *Widerspruchsherleitung*, esto es, ‘derivación de una contradicción’. El término es correctísimo, pero el equivalente castellano es demasiado largo para usarlo a cada momento. La versión inglesa dice ‘contradictive derivation’, lo que sugiere falsamente que habría un vicio lógico en la derivación misma.

12 Para que esta definición sea aplicable hay que suponer escritos todos los paréntesis que convencionalmente omitimos; por ejemplo, ‘A ∧ B’ es propiamente ‘(A ∧ B)’. El alcance de los conectivos está definido en el Apéndice IX.I.

Un *hilo* en la derivación **Δ** es una lista de secuentes $\langle\sigma_1, \ldots, \sigma_n\rangle$ que figuran en **Δ**, cada uno de los cuales precede inmediatamente al siguiente (en otras palabras, si $1 \leq k < n$, $\sigma_k$ aparece en **Δ** como premisa de una inferencia cuya conclusión es $\sigma_{k+1}$). Diré que el hilo $\langle\sigma_1, \ldots, \sigma_n\rangle$ es un *hilo completo* en **Δ** si $\sigma_1$ es un secuente inicial en **Δ** y $\sigma_n$ es el secuente final de **Δ**.[13] Consideremos ahora todos los hilos en la derivación **Δ** que (i) contienen el secuente final y (ii) no contienen ninguna inferencia con conectivo. La unión de todos estos hilos constituye un árbol de secuentes que es una parte —generalmente, una parte propia— de la derivación entera. Lo llamaremos, con Gentzen, el *trozo final* (*Endstück*) de la derivación **Δ**.

Supongamos que **Δ** es una derivación fatal. Con este supuesto, construiremos otra derivación fatal **Δ*** diseñada de modo que preceda a **Δ** en el orden que se definirá en la Sección 2.12.3 (en cierto modo no enteramente obvio, **Δ*** es más simple que **Δ**).

*Eliminación de las variables libres ociosas*. Podemos suponer, sin pérdida de generalidad, que **Δ** no contiene ninguna variable libre que no sea la variable propia de una inferencia. En efecto, las variables libres que no tienen éste carácter no contribuyen en nada a una derivación y pueden todas reemplazarse, digamos, por la constante 1 sin que ello afecte la corrección de las inferencias. Podemos asimismo suponer que, si $\mathfrak{a}$ es la variable propia de una inferencia $\mathscr{I}$ en **Δ**, $\mathfrak{a}$ no es la variable propia de ninguna inferencia situada sobre $\mathscr{I}$. En efecto, si $\mathfrak{a}$ es también la variable libre de otra inferencia $\mathscr{I}^\circ$ situada sobre $\mathscr{I}$, se la puede reemplazar en todas las posiciones que ocupa sobre la conclusión de $\mathscr{I}^\circ$ por una variable nueva —que no figure en toda la derivación— sin afectar la corrección de las inferencias. (Recuérdese que la variable propia de una inferencia no puede figurar en la conclusión de esa inferencia).

*Eliminación de las inferencias inductivas en el trozo final*. Sea **Z** el trozo final de la derivación **Δ**. Por definición, **Z** sólo puede contener inferencias estructurales o inductivas. Supongamos que contiene una o más inferencias inductivas. Elijamos una bajo la cual no haya otra inferencia de esa clase.

13 Gentzen (1938, p. 24) define ‘hilo’ (*Faden*) en los términos que he utilizado para definir un ‘hilo completo’, pero luego emplea el término ‘hilo’ como si lo estuviera entendiendo de acuerdo con mi definición.

Ella tiene la forma

$$\frac{\mathfrak{F}(\mathfrak{a}/\mathfrak{x}), \Gamma \rightarrow \Theta, \mathfrak{F}(\mathfrak{a}'/\mathfrak{x})}{\mathfrak{F}(1/\mathfrak{x}), \Gamma \rightarrow \Theta, \mathfrak{F}(\mathfrak{n}/\mathfrak{x})}$$

donde $\mathfrak{n}$ es un término numérico. Sea $\mathfrak{n} \neq 1$ (el caso $\mathfrak{n} = 1$ se examina en la nota 16). La conclusión no puede contener *ninguna* variable libre, porque bajo ella no hay ninguna inferencia con variables propias.[14] Para facilitar las referencias, llamo **E** al trozo de $\mathbf{\Delta}$ formado por esta inferencia inductiva y todo lo que hay sobre ella. Obviamente, **E** constituye por sí misma una derivación cuyo secuente final es $\mathfrak{F}(1/\mathfrak{x}), \Gamma \rightarrow \Theta, \mathfrak{F}(\mathfrak{n}/\mathfrak{x})$. Ahora bien, este secuente se puede derivar mediante *inferencias estructurales* de las premisas $\mathfrak{F}(1/\mathfrak{x}), \Gamma \rightarrow \Theta, \mathfrak{F}(1'/\mathfrak{x})$; $\mathfrak{F}(1'/\mathfrak{x}), \Gamma \rightarrow \Theta, \mathfrak{F}(1''/\mathfrak{x})$; $\mathfrak{F}(1''/\mathfrak{x}), \Gamma \rightarrow \Theta, \mathfrak{F}(1'''/\mathfrak{x})$; . . . ; $\mathfrak{F}(\mathfrak{n}–1/\mathfrak{x}), \Gamma \rightarrow \Theta, \mathfrak{F}(\mathfrak{n}/\mathfrak{x})$ (donde $\mathfrak{n}–1$ representa el término numérico que tiene un signo $'$ menos que el representado por $\mathfrak{n}$), como indica el siguiente esquema parcial,

$$\begin{array}{c}
\mathfrak{F}(1/\mathfrak{x}), \Gamma \rightarrow \Theta, \mathfrak{F}(1'/\mathfrak{x}) \qquad \mathfrak{F}(1'/\mathfrak{x}), \Gamma \rightarrow \Theta, \mathfrak{F}(1''/\mathfrak{x}) \\
\hline
\mathfrak{F}(1'/\mathfrak{x}), \Gamma, \Gamma \rightarrow \Theta, \Theta, \mathfrak{F}(1''/\mathfrak{x}) \quad \text{(corte)} \\
\hline\hline
\mathfrak{F}(1/\mathfrak{x}), \Gamma \rightarrow \Theta, \mathfrak{F}(1''/\mathfrak{x}) \qquad \mathfrak{F}(1''/\mathfrak{x}), \Gamma \rightarrow \Theta, \mathfrak{F}(1'''/\mathfrak{x}) \\
\hline\hline
\mathfrak{F}(1/\mathfrak{x}), \Gamma, \Gamma \rightarrow \Theta, \Theta, \mathfrak{F}(1'''/\mathfrak{x}) \quad \text{(corte)} \\
\vdots \\
\hline
\mathfrak{F}(1/\mathfrak{x}), \Gamma \rightarrow \Theta, \mathfrak{F}(\mathfrak{n}/\mathfrak{x}) \quad \text{(corte)}
\end{array}$$

donde las rayas dobles representan contracciones y permutaciones y la construcción continúa en forma similar en el trecho entre los últimos dos cortes.

[14] En el trozo final no hay inferencias con variables propias de los tipos $\forall$ o $\exists$, ni inferencias tipo IM bajo la elegida.

Llamaré *premisas iniciales* del esquema a los secuentes de la forma

$$\mathfrak{F}(\mathfrak{m}/\mathfrak{x}),\ \Gamma \twoheadrightarrow \Theta,\ \mathfrak{F}(\mathfrak{m}'/\mathfrak{x})$$

que en él son premisas sin ser conclusiones. En la derivación $\mathbf{\Delta}$, la premisa $\mathfrak{F}(\mathfrak{a}/\mathfrak{x}), \Gamma \twoheadrightarrow \Theta, \mathfrak{F}(\mathfrak{a}'/\mathfrak{x})$ de la última inferencia inductiva es el secuente final de una derivación **H** (**H** es igual al trozo de $\mathbf{\Delta}$ que llamé **E**, quitando la última raya horizontal y el secuente bajo ella). En **H** la variable libre $\mathfrak{a}$ no es la variable propia de ninguna inferencia.[15] Por lo tanto, se la puede reemplazar uniformemente por un dado término numérico, sin afectar la corrección de las inferencias. Sea $\mathbf{H}(\mathfrak{m}/\mathfrak{a})$ la derivación obtenida al reemplazar la variable $\mathfrak{a}$ por el término numérico $\mathfrak{m}$ en todas las posiciones libres de $\mathfrak{a}$ en **H**. Reemplacemos en el último esquema, la premisa inicial $\mathfrak{F}(1/\mathfrak{x}), \Gamma \twoheadrightarrow \Theta, \mathfrak{F}(1'/\mathfrak{x})$ por $\mathbf{H}(1/\mathfrak{a})$, la premisa inicial $\mathfrak{F}(1'/\mathfrak{x}), \Gamma \twoheadrightarrow \Theta, \mathfrak{F}(1''/\mathfrak{x})$ por $\mathbf{H}(1'/\mathfrak{a})$ y, en general, cada premisa inicial de la forma $\mathfrak{F}(\mathfrak{m}/\mathfrak{x}), \Gamma \twoheadrightarrow \Theta, \mathfrak{F}(\mathfrak{m}'/\mathfrak{x})$ por $\mathbf{H}(\mathfrak{m}/\mathfrak{a})$. El producto de este reemplazo es una derivación que llamaré $\mathbf{E}_1$. Si reemplazamos en $\mathbf{\Delta}$ el trozo **E** por $\mathbf{E}_1$ obtenemos una derivación fatal $\mathbf{\Delta}_1$ que contiene una inferencia inductiva menos que $\mathbf{\Delta}$.[16] Repitiendo este procedimiento un número finito de veces se obtiene una derivación fatal $\mathbf{\Delta}^\circ$ que no contiene ninguna inferencia inductiva en su trozo final.

*Alianza de fórmulas en una derivación.* Nuestro próximo paso será mostrar que, si $\mathbf{\Delta}^\circ$ es una derivación fatal que sólo contiene inferencias estructurales en su trozo final, hay una derivación fatal $\mathbf{\Delta}^\star$ que no contiene *debilitamientos* ni *secuentes básicos lógicos* (SBLs) en su trozo final. Pero antes definiré un concepto que luego nos será muy útil. Es una relación que agrupa fórmulas iguales (*copias* de una misma fórmula) en lugares diferentes en una derivación; esto es, una relación entre fórmulas *situadas* (marcadas, si se quiere, con un índice de su posición). En una inferencia estructural hay, por re-

15 De otro modo, $\mathfrak{a}$ sería la variable propia de la inferencia elegida y de otra inferencia que la precede. Esto contradiría nuestra suposición inicial sobre la unicidad de las variables propias.

16 Consideremos ahora el caso —que arriba dejé de lado— en que el término numérico $\mathfrak{n}$ en la conclusión de la última inferencia inductiva de $\mathbf{\Delta}$ es la constante 1. Entonces, esa conclusión puede derivarse del SBL $\mathfrak{F}(1/\mathfrak{x}) \twoheadrightarrow \mathfrak{F}(1/\mathfrak{x})$, por debilitamientos y permutaciones. Llamemos $\mathbf{E}_0$ a esta derivación trivial. Reemplazando el trozo **E** por *esta* derivación $\mathbf{E}_0$ se obtiene una derivación fatal que contiene una inferencia inductiva menos que la derivación $\mathbf{\Delta}$.

gla general, grupos de fórmulas situadas que consideramos homólogas; por ejemplo, las tres fórmulas representadas por la letra $\mathfrak{D}$ en los esquemas de la contracción, las dos representadas por $\mathfrak{D}$ en los esquemas de permutación y corte, las dos representadas por $\mathfrak{E}$ en los esquemas de permutación; pero también las fórmulas que ocupan la misma posición (primera, segunda, . . . , enésima) en las dos copias de la lista de fórmulas representada en un esquema por cierta mayúscula griega. Diré que dos fórmulas iguales que ocupan lugares diferentes en una derivación están *aliadas* si figuran en la misma inferencia y son homólogas en el sentido indicado. Sea $\mathfrak{F}$ una fórmula situada en el trozo final **Z** de una derivación **Δ**. La *alianza* de $\mathfrak{F}$ en **Δ** es el conjunto $\mathscr{A}_{\mathfrak{F}}$ de fórmulas situadas en **Z** caracterizado así: (i) $\mathfrak{F} \in \mathscr{A}_{\mathfrak{F}}$; (ii) si $\mathfrak{F}_1 \in \mathscr{A}_{\mathfrak{F}}$ y $\mathfrak{F}_2$ está aliada a $\mathfrak{F}_1$, $\mathfrak{F}_2 \in \mathscr{A}_{\mathfrak{F}}$.[17] Ahora bien, si **Δ** es una derivación fatal, es claro que $\mathscr{A}_{\mathfrak{F}}$ contiene las dos fórmulas cortadas en un corte. En efecto, toda fórmula de $\mathscr{A}_{\mathfrak{F}}$ que no sea una de las fórmulas cortadas en un corte está aliada a una fórmula del secuente que sigue inmediatamente al suyo. Como el secuente final de **Δ** no contiene fórmulas, **Z** tiene que incluir un corte —y sólo uno[18]— cuyas fórmulas cortadas sean copias de $\mathfrak{F}$ y pertenezcan a $\mathscr{A}_{\mathfrak{F}}$. Lo llamo *el corte propio* de la alianza $\mathscr{A}_{\mathfrak{F}}$. Todas las fórmulas de la alianza $\mathscr{A}_{\mathfrak{F}}$ están situadas sobre la conclusión de su corte propio. Distinguimos el *lado izquierdo* de $\mathscr{A}_{\mathfrak{F}}$, formado por copias de $\mathfrak{F}$ situadas en hilos que terminan en la primera premisa del corte propio, y el *lado derecho* de $\mathscr{A}_{\mathfrak{F}}$, formado por copias de $\mathfrak{F}$ situadas en hilos que terminan en la segunda premisa del corte propio. Se advertirá que todas las fórmulas del lado izquierdo son posfórmulas de sus respectivos secuentes, y que todas las del derecho son prefórmulas. Cada lado de la alianza tiene la estructura de un árbol cuya raíz es una de las dos fórmulas cortadas en el corte propio. Éste se ramifica dondequiera haya copias de $\mathfrak{F}$ aliadas en la premisa de una contracción. Si un elemento de un lado es la fórmula agregada en la conclusión de un debilitamiento o pertenece a un secuente inicial de la derivación, decimos que es una *fórmula inicial* de ese lado y de la alianza $\mathscr{A}_{\mathfrak{F}}$.

17 ‘Aliada’ y ‘alianza’ corresponden, respectivamente, a los términos ‘gebunden’ y ‘Bund’, empleados por Gentzen. La traducción inglesa dice ‘clustered’ y ‘cluster’.

18 Si hubiera dos, uno precedería al otro; pero las fórmulas cortadas en un corte no pueden pertenecer a una alianza que incluya fórmulas situadas bajo las premisas de ese corte.

*Eliminación de los debilitamientos y los secuentes básicos lógicos en el trozo final*. Ahora procederemos a eliminar los debilitamientos y los SBLs del trozo final $\mathbf{Z}^\circ$ de la derivación fatal $\mathbf{\Delta}^\circ$.

Sea $\mathfrak{D}$ un debilitamiento que figura en $\mathbf{Z}^\circ$, sobre el cual no hay ningún otro debilitamiento en $\mathbf{Z}^\circ$. Sea $\mathfrak{F}$ la fórmula agregada en la conclusión de $\mathfrak{D}$. Para formar una nueva derivación fatal que no contenga la inferencia $\mathfrak{D}$ simplemente eliminamos la conclusión de $\mathfrak{D}$ y todas las copias de $\mathfrak{F}$ que componen la alianza $\mathcal{A}_{\mathfrak{F}}$. Dos casos merecen atención. (i) Puede haber dos copias de $\mathfrak{F}$ aliadas en la premisa de una contracción; al borrarlas junto con su aliada en la conclusión, la nueva conclusión resulta ser una copia de la nueva premisa y, por ende, se la puede eliminar completamente (la contracción desaparece). (ii) El corte propio de $\mathcal{A}_{\mathfrak{F}}$ contiene una copia de $\mathfrak{F}$ que no está al mismo lado que $\mathfrak{F}$; esa copia se elimina junto con el secuente a que pertenece y todos los secuentes situados sobre él; la otra copia de $\mathfrak{F}$ en el corte propio de $\mathcal{A}_{\mathfrak{F}}$ se borra simplemente (como las demás en su lado); la conclusión del corte se infiere de la premisa sobreviviente por debilitamientos y permutaciones, como se verá comparando los esquemas siguientes.

$$\frac{\Gamma \rightarrow \Theta, \mathfrak{D} \qquad \Delta, \mathfrak{D} \rightarrow \Lambda}{\Gamma, \Delta \rightarrow \Theta, \Lambda} \qquad\qquad \frac{\Gamma \rightarrow \Theta}{\overline{\Gamma, \Delta \rightarrow \Theta, \Lambda}}$$

Reflexionando sobre los esquemas de las inferencias estructurales el lector se persuadirá de que el producto de la transformación a que $\mathbf{\Delta}^\circ$ ha sido sometida también es una derivación fatal, que contiene un debilitamiento menos que $\mathbf{\Delta}^\circ$. Repitiendo el mismo proceso un número finito de veces, se obtiene una derivación fatal $\mathbf{\Delta}^{\circ\circ}$ que no contiene debilitamientos en su trozo final $\mathbf{Z}^{\circ\circ}$.

Si $\mathbf{Z}^\circ$ contiene uno o más SBLs, elijamos uno que represento con $\mathfrak{A} \rightarrow \mathfrak{A}$. Como el secuente final no contiene fórmulas, $\mathfrak{A} \rightarrow \mathfrak{A}$ tiene que ser la premisa de una inferencia. No puede serlo de una permutación ni de una contracción. Por lo tanto, tiene que ser una de las premisas de un corte. Entonces la conclusión de ese corte será igual a la otra premisa, como se ve en los esquemas siguientes.

$$\frac{\mathfrak{A} \rightarrowtail \mathfrak{A} \qquad \mathfrak{A}, \Delta \rightarrowtail \Lambda}{\mathfrak{A}, \Delta \rightarrowtail \Lambda} \qquad \frac{\Gamma \rightarrowtail \Theta \qquad \mathfrak{A} \rightarrowtail \mathfrak{A}}{\Gamma \rightarrowtail \Theta, \mathfrak{A}}$$

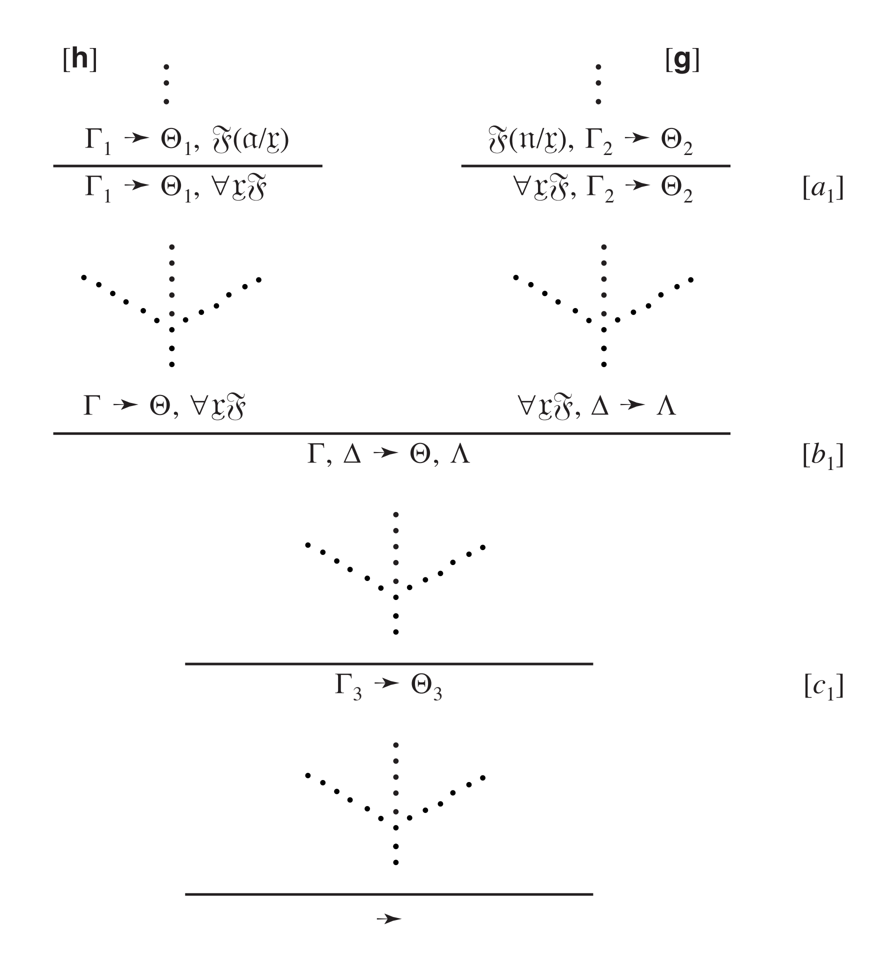
Es posible, pues, eliminar el corte, borrando el SBL $\mathfrak{A} \rightarrowtail \mathfrak{A}$ e identificando la conclusión con la otra premisa, sin detrimento de las demás inferencias. Repitiendo este proceso se eliminan en un número finito de pasos todos los SBLs contenidos en el trozo final de la derivación $\mathbf{\Delta}^{\circ\circ}$.

*Reducción del conectivo*. Hemos demostrado que si existe una derivación fatal $\mathbf{\Delta}$, también hay una derivación fatal $\mathbf{\Delta}^{\star}$ cuyo trozo final $\mathbf{Z}^{\star}$ no contiene variables libres, inferencias inductivas, debilitamientos o secuentes básicos lógicos. Si $\mathbf{\Delta}^{\star}$ no contiene ninguna inferencia con conectivo, $\mathbf{\Delta}^{\star}$ coincide con su trozo final. Las características de $\mathbf{\Delta}^{\star}$ implican, en tal caso, que todos sus secuentes iniciales son secuentes básicos *matemáticos* (SBMs). Estos constan exclusivamente de fórmulas elementales. Como no contienen variables libres —puesto que $\mathbf{\Delta}^{\star}$ no las contiene— dichos secuentes expresan verdades aritméticas decidibles. Evidentemente, el secuente ‘ $\rightarrowtail$ ’ no se puede derivar de tales secuentes mediante cortes, contracciones y permutaciones. Por lo tanto, si $\mathbf{\Delta}^{\star}$ es una derivación fatal, $\mathbf{\Delta}^{\star}$ contiene al menos una inferencia con conectivo. Voy a mostrar que en tal caso existe una derivación fatal $\mathbf{\Delta}^{\dagger}$ que precede a $\mathbf{\Delta}^{\star}$ en el orden que definiremos en la Sección 2.12.3 (como allí se verá). Llamaré a la transformación de $\mathbf{\Delta}^{\star}$ en $\mathbf{\Delta}^{\dagger}$ la *reducción del conectivo* —ya que Gentzen la llama *Verknüpfungsreduktion*— pero ella no envuelve la *eliminación* de un conectivo; la preposición ‘de’ significa aquí ‘concerniente a’.

El trozo final $\mathbf{Z}^{\star}$ de la derivación $\mathbf{\Delta}^{\star}$ está formado por hilos que convergen en el secuente final ‘ $\rightarrowtail$ ’. Algunos de esos hilos pueden empezar —por arriba— con SBMs que no contienen variables libres, pero al menos uno de ellos empieza con la conclusión de una inferencia con conectivo, cuya premisa o premisas quedan, por definición, fuera de $\mathbf{Z}^{\star}$. Si hay varios hilos de esta clase, asignamos un gödel a cada uno y elegimos el de numeración más baja, que llamaré $\mathsf{h}$. Sea $\mathfrak{H}$ la fórmula principal de la inferencia con conectivo cuya conclusión constituye el extremo superior del hilo elegido. Diré que $\mathfrak{H}$ es la *fórmula crítica* de la reducción. Consideremos la alianza $\mathscr{A}_{\mathfrak{H}}$ que $\mathfrak{H}$ forma dentro de $\mathbf{Z}^{\star}$ con otras fórmulas iguales a ella. En el corte propio de $\mathscr{A}_{\mathfrak{H}}$ las fórmulas cortadas son copias de $\mathfrak{H}$. Por lo tanto, la premi-

sa de dicho corte que no está en el hilo **h**, pertenece a otro hilo **g** cuyo extremo superior también es la conclusión de una inferencia con conectivo. (De otro modo, dicha premisa pertenecería sólo a uno o más hilos completos contenidos enteramente en el trozo final $\mathbf{Z}^{\star}$, que no contiene debilitamientos ni SBLs, y por ende no podría contener una fórmula con un conectivo). Para fijar ideas, supongamos que $\forall$ es el conectivo principal de la fórmula crítica de la reducción. Entonces la derivación fatal $\mathbf{\Delta}^{\star}$ contiene el trozo representado esquemáticamente en el Cuadro 1:

CUADRO 1



Las líneas punteadas sobre los primeros secuentes anotados representan las ramas de $\mathbf{\Delta}^{\star}$ que confluyen en los hilos **h** y **g**, respectivamente. Cualquier otra línea punteada vertical representa la continuación de **h** y **g** entre los dos secuentes que enmarcan esa línea. Las líneas punteadas inclinadas que salen de una misma vertical representan los hilos —0, 1, 2 o más— que desembocan en el representado por ésta, entre los dos secuentes que la enmarcan. El término $\mathfrak{n}$ en la premisa de la inferencia con conectivos al lado derecho tiene que ser un término numérico, ya que esa no es una inferencia con variable propia y no hay otras inferencias con variable propia bajo ella.[20] He escrito [$a_1$] frente a las dos conclusiones de inferencias con conectivo con que empiezan los hilos **h** y **g**; [$b_1$] frente al corte propio de la alianza de la fórmula $\forall\mathfrak{x}\mathfrak{F}$, y [$c_1$] frente al primer secuente bajo dicho corte cuya altura es menor que la altura de las premisas de ese corte (tiene que haber un secuente así, puesto que el secuente final tiene altura 0).[21] Este secuente puede ser la conclusión de ese corte, en cuyo caso las líneas marcadas con [$b_1$] y [$c_1$] coinciden. También puede ocurrir que las premisas del corte sean las conclusiones marcadas con [$a_1$] y que su conclusión sea el secuente final ‘ ➛ ’. Estas situaciones especiales simplifican la reducción del conectivo sin afectarla.

La reducción del conectivo $\forall$ en la posición indicada transforma la derivación $\mathbf{\Delta}^{\star}$ en la derivación $\mathbf{\Delta}^{\dagger}$ presentada esquemáticamente en el Cuadro 2 (p. 456). Las líneas punteadas marcadas con [**h**] y [**g**] representan sendas copias de los árboles marcados del mismo modo en el Cuadro 1. Se ha invertido el orden para acomodar dos árboles nuevos marcados [**h′**] y [**g′**] que se combinan en cortes con la continuación de [**h**] y [**g**], como se indica en la línea [$c_2$]. Hasta la línea [$a_2$], [**g′**] es simplemente una copia de [**g**] y [**h′**] es el resultado de reemplazar en [**h**] la variable $\mathfrak{a}$, en todas sus posiciones libres, por el término numérico $\mathfrak{n}$. Bajo la línea [$a_2$], [**h′**] y [**g′**] se continúan con permutaciones y un debilitamiento para obtener las conclusiones que ocupan la línea [$b_2$]. La línea [$d_2$] contiene los primeros secuentes cuya altu-

20 Según acordamos arriba, la derivación entera no contiene variables libres que no sean variables propias de una inferencia. En virtud de la eliminación de las inferencias inductivas en el trozo final, éste no contiene ninguna inferencia con variable propia.

21 Recuérdese que la *altura* de un secuente en una derivación es el más alto grado poseído por un corte o inferencia inductiva cuya conclusión está bajo ese secuente.

ra es inferior a la de las premisas del corte respectivo en la línea [$c_2$] ([$d_2$] corresponde pues a [$c_1$] en el Cuadro 1). Los secuentes de la línea [$e_2$] se derivan por cero, una o más permutaciones de sus respectivos predecesores en la línea [$d_2$]. La línea [$f_2$] contiene la conclusión de lo que llamaremos con Gentzen *el nuevo corte*. La línea [$g_2$] se deriva de [$f_2$] mediante contracciones y permutaciones. Las rayas horizontales dobles representan, pues, cero, una o más inferencias estructurales triviales, según haga falta en cada caso.[22] A la luz de estas indicaciones es fácil ver que $\boldsymbol{\Delta}^\dagger$ es una derivación genuina. Aunque, a primera vista, $\boldsymbol{\Delta}^\dagger$ parece más compleja que $\boldsymbol{\Delta}^\star$, una mirada más atenta descubre sutiles simplificaciones. El corte propio de la alianza de $\forall\mathfrak{x}\mathfrak{F}\mathfrak{x}$ en $\boldsymbol{\Delta}^\star$, marcado con [$b_1$], es en cierto sentido homólogo a los cortes marcados con [$c_2$] en $\boldsymbol{\Delta}^\dagger$; pero sobre cada uno de estos cortes hay *una* inferencia con conectivo *menos* que sobre el corte propio de la alianza de $\forall\mathfrak{x}\mathfrak{F}\mathfrak{x}$ en $\boldsymbol{\Delta}^\star$. En otro sentido, al corte propio de la alianza de $\forall\mathfrak{x}\mathfrak{F}\mathfrak{x}$ en $\boldsymbol{\Delta}^\star$ le corresponde en $\boldsymbol{\Delta}^\dagger$ el nuevo corte (entre las líneas [$e_2$] y [$f_2$]); pero el *grado* de éste —esto es, el número de conectivos en las fórmulas cortadas— es *menor* que el de aquél. Como pronto veremos, el orden que Gentzen establece entre las derivaciones está pensado para aprovechar estas diferencias.

Si el conectivo principal de la fórmula crítica de la reducción no es $\forall$ hay que modificar algunos detalles de la transformación. Si es $\wedge$, la fórmula crítica tiene la forma $\mathfrak{A} \wedge \mathfrak{B}$. Entonces, en el esquema de la derivación $\boldsymbol{\Delta}^\star$ la inferencia con conectivo al lado izquierdo tendrá la forma

$$\frac{\Gamma_1 \rightarrow \Theta_1, \mathfrak{A} \qquad\qquad \Gamma_1 \rightarrow \Theta_1, \mathfrak{B}}{\Gamma_1 \rightarrow \Theta_1, \mathfrak{A} \wedge \mathfrak{B}} \qquad [a]$$

y la inferencia con conectivo al lado derecho tendrá una de las formas

$$\frac{\mathfrak{A}, \Gamma_2 \rightarrow \Theta_2}{\mathfrak{A} \wedge \mathfrak{B}, \Gamma_2 \rightarrow \Theta_2} \qquad \frac{\mathfrak{B}, \Gamma_2 \rightarrow \Theta_2}{\mathfrak{A} \wedge \mathfrak{B}, \Gamma_2 \rightarrow \Theta_2} \qquad [a]$$

22 Puede ocurrir que el secuente que precede inmediatamente a un par de rayas tenga ya la forma del secuente que inmediatamente las sigue; en tales casos, se sobreentiende que la derivación no contiene dichas inferencias y que hay un solo secuente donde el esquema presenta dos.

Bajo la línea [*a*], hallaremos también $\mathfrak{A} \wedge \mathfrak{B}$ en lugar de $\forall\mathfrak{x}\mathfrak{F}$ y una de las fórmulas $\mathfrak{A}$ o $\mathfrak{B}$ en vez de $\mathfrak{F}(\mathfrak{n}/\mathfrak{x})$. La derivación $\mathbf{\Delta}^{\dagger}$ se ajusta facilmente a estos cambios (ensáyelo el lector). En particular, el árbol marcado [**h′**] contiene sobre la primera raya horizontal solamente una derivación del secuente $\mathfrak{A}, \Gamma_2 \succ \Theta_2$ (respectivamente, $\mathfrak{B}, \Gamma_2 \succ \Theta_2$).

Si el conectivo principal de la fórmula crítica de la reducción es $\neg$, las dos inferencias con conectivo en el esquema de la derivación $\mathbf{\Delta}^{\star}$ tendrán la forma

$$\frac{\mathfrak{A}, \Gamma_1 \succ \Theta_1}{\Gamma_1 \succ \Theta_1, \neg\mathfrak{A}} \qquad \frac{\Gamma_2 \succ \Theta_2, \mathfrak{A}}{\neg\mathfrak{A}, \Gamma_2 \succ \Theta_2} \qquad [a]$$

En la derivación $\mathbf{\Delta}^{\dagger}$ los árboles designados con [**h′**] y [**g**] deben ir a la derecha de los designados con [**h**] y [**g′**]. Las partes superiores de [**h′**] y [**g′**] tienen la forma

$$\begin{array}{ccc} [\mathbf{g'}] \ \vdots & \quad & \vdots \ [\mathbf{h'}] \\ \Gamma_2 \succ \Theta_2, \mathfrak{A} & & \mathfrak{A}, \Gamma_1 \succ \Theta_1 \\ \hline\hline \neg\mathfrak{A}, \Gamma_2 \succ \Theta_2, \mathfrak{A} & & \mathfrak{A}, \Gamma_1 \succ \Theta_1, \neg\mathfrak{A} \end{array} \qquad \begin{array}{c} \\ [a] \\ \\ [b] \end{array}$$

El lector debe hacer los cambios correspondientes en el resto del esquema que representa a $\mathbf{\Delta}^{\dagger}$ y verificar que el resultado es una derivación genuina. Dejo también a su cuidado el caso de los dos conectivos restantes, que resolverá sin dificultad aprovechando las relaciones de dualidad entre $\vee$ y $\wedge$ y entre $\forall$ y $\exists$. (Cambie las prefórmulas en posfórmulas y viceversa y ponga a la derecha el árbol de la izquierda y a la izquierda el de la derecha).

### *2.12.3 Orden de las derivaciones e inducción transfinita*

Gentzen ordena las derivaciones asignándole a cada una un elemento de una colección bien ordenada de objetos definidos por él. Aunque los llama 'números ordinales' (*Ordnungszahlen*) y los designa con símbolos tomados de Cantor, Gentzen advierte que esos símbolos y las operaciones con que los combina "han de entenderse para nuestros propósitos de un modo enteramente *formal*, sin que haya que atribuirles ningún *sentido*, como que ω es 'un número infinito' y el signo '+' representa la 'adición'" (1938, p. 38). Luego diré algo sobre el sentido y los méritos de esta advertencia. Pero atendamos primero a la definición de los ordinales de Gentzen.

Gentzen define recursivamente una secuencia infinita $\mathfrak{S}_0, \mathfrak{S}_1, \mathfrak{S}_2, \ldots$ de "sistemas", y relaciones de igualdad (=) y de orden (>) entre los objetos que los forman. El sistema $\mathfrak{S}_0$ consta de un solo objeto, llamado 0. En $\mathfrak{S}_0$, las dos relaciones mencionadas se definen por las condiciones: (i) 0 = 0 y (ii) es falso que 0 > 0. Supongamos ahora que ya se ha definido la composición del sistema $\mathfrak{S}_n$, así como las relaciones de igualdad y orden en $\mathfrak{S}_n$ ($n \in \mathbb{N}$).[23] Entonces, el sistema $\mathfrak{S}_{n+1}$ comprende (i) el 0 y (ii) todos los objetos de la forma

$$\omega^{\alpha_1} + \omega^{\alpha_2} + \ldots + \omega^{\alpha_m} \qquad (*)$$

donde $m \in \mathbb{N}$, $\alpha_1, \alpha_2, \ldots, \alpha_m$ pertenecen a $\mathfrak{S}_n$ y $\alpha_1 \geq \alpha_2 \geq \ldots \geq \alpha_m$. Las relaciones de igualdad y orden en $\mathfrak{S}_{n+1}$ se definen así: Sean $\alpha = \omega^{\alpha_1} + \omega^{\alpha_2} + \ldots + \omega^{\alpha_p}$ y $\beta = \omega^{\beta_1} + \omega^{\beta_2} + \ldots + \omega^{\beta_q}$ elementos de $\mathfrak{S}_{n+1}$ ($p \leq q$); entonces

(*a*) $\alpha = \beta$ si y sólo si $p = q$ y $\alpha_k = \beta_k$ ($1 \leq k \leq q$);
(*b*) $\alpha > \beta$ si $\alpha_i = \beta_i$ para todo $i < k \leq q$, $\alpha_k \neq \beta$ y $\alpha_k > \beta_k$, o si $\alpha_k = \beta_k$ para todo $k \leq q$ y $q < p$;
(*c*) $\alpha > 0$, a menos que $\alpha = 0$.

[23] Gentzen escribe: "Die Zahlen des Systems $\mathfrak{S}_\rho$ (ρ sei eine natürliche Zahl oder 0) seien bereits definiert, ebenso = und <-Beziehung zwischen diesen" (1938, p. 38). Tendré ocasión de referirme a este pasaje al final de esta sección.

Como de costumbre, si $\alpha > \beta$ (en palabras: $\alpha$ *es mayor que* $\beta$), decimos también que $\beta < \alpha$ (en palabras: $\beta$ *es menor que* $\alpha$). Es claro que si $\alpha$ y $\beta$ pertenecen a $\mathfrak{S}_{n+1}$, se cumple una y solo una de las tres condiciones siguientes: $\alpha > \beta$, $\alpha = \beta$ o $\alpha < \beta$.

Los ordinales de Gentzen son los elementos de los sistemas $\mathfrak{S}_0$, $\mathfrak{S}_1$, $\mathfrak{S}_2, \ldots$ Como el lector puede fácilmente verificar, si $n$ y $m$ son dos números naturales tales que $n$ es menor que $m$, $\mathfrak{S}_n$ está incluido en $\mathfrak{S}_m$. La inyección canónica $\mathfrak{S}_n \to \mathfrak{S}_m$, $x \mapsto x$ obviamente preserva las relaciones de igualdad y orden. Es obvio asimismo que, dada una expresión cualquiera, es posible decidir efectivamente si posee o no la forma (*) y, por lo tanto, si es o no un ordinal. Dadas dos expresiones de la forma (*) es posible decidir efectivamente si son o no iguales, y en caso que no lo sean, cuál de las dos es mayor. (También, por cierto, si una de las expresiones dadas es el 0). Por lo tanto, todos estos conceptos son "finitistas" (*"finit"*) en el sentido de Hilbert, concluye Gentzen (1938, p. 38). Con todo, la misma notación adoptada conduce inevitablemente a establecer una correspondencia biunívoca entre los ordinales de Gentzen y un segmento de los ordinales de Cantor, como sigue: $0 \mapsto 0$, $\omega^0 \mapsto 1$, $\omega^0 + \omega^0 \mapsto 2, \ldots$ , $\omega^1 \mapsto \omega$, $\omega^1 + \omega^0 \mapsto \omega + 1, \ldots$ En suma, a cada ordinal de Gentzen le corresponde de manera exclusiva su homónimo en el sistema de Cantor. Según Gentzen, esta correspondencia se introduce —así como la nomenclatura que la sugiere— sólo para facilitar la comprensión a los lectores familiarizados con la teoría de conjuntos. Pero es mucho más profunda que las analogías y homonimias didácticas habituales. La correspondencia es un isomorfismo de órdenes, puesto que la relación > entre los ordinales de Gentzen se definió, como es obvio, en armonía con la relación 'mayor que' entre sus homónimos cantorianos. En virtud del isomorfismo, cada sistema gentzeniano $\mathfrak{S}_k$ ($k \geq 0$) corresponde, a un conjunto de ordinales cantorianos que también llamaré $\mathfrak{S}_k$. Tales conjuntos contienen los ordinales que se indican a continuación:

$$\mathfrak{S}_0 = \{0\}$$
$$\mathfrak{S}_1 = \mathfrak{S}_0 \cup \{1, 2, \ldots\} = \{x\text{: } x < \omega\}$$
$$\mathfrak{S}_2 = \mathfrak{S}_1 \cup \{\omega, \omega+1, \omega+2, \ldots\} = \{x\text{: } x < \omega^\omega\}$$
$$\mathfrak{S}_3 = \{x\text{: } x < \omega^{\omega^\omega}\}$$
$$\mathfrak{S}_4 = \{x\text{: } x < \omega^{\omega^{\omega^\omega}}\}$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

Cantor llamó $\varepsilon_0$ al ordinal más pequeño que satisface la ecuación $\omega^\xi = \xi$. Es claro, pues, que el sistema ordenado que forman todos los ordinales de Gentzen es isomórfico a $\varepsilon_0$. El isomorfismo permite transferir a los ordinales de Gentzen las operaciones de adición, multiplicación y exponenciación transfinitas definidas en el Capítulo 1.5. En tal caso, el signo +, no obstante la advertencia de Gentzen, funciona precisamente como signo de adición. En su exposición, Gentzen aduce la multiplicación y la exponenciación, pero advierte enfáticamente que estos conceptos "sirven, por cierto, sólo para aclarar las cosas, ya que de ninguna manera están definidos para nosotros" (1938, p. 40).[24] Por otra parte, introduce formalmente una operación algebraica binaria entre sus ordinales, como sigue. Sean $\alpha = \omega^{\alpha} + \omega^{\alpha_2} + \ldots + \omega^{\alpha_p}$ y $\beta = \omega^{\beta_1} + \omega^{\beta_2} + \ldots + \omega^{\beta_q}$ dos ordinales de Gentzen. Entonces la *suma natural* $\alpha \# \beta$ es el ordinal que se obtiene escribiendo $p + q$ copias de $\omega$ separadas por el signo + y asignándoles los "exponentes" $\alpha_1, \ldots, \alpha_p, \beta_1, \ldots, \beta_q$, ordenados de mayor a menor. Evidentemente, la suma natural es una operación conmutativa y asociativa. Es fácil probar que $(\alpha \# \beta) > \alpha$ y que si $\alpha' > \alpha$, $(\alpha' \# \beta) > (\alpha \# \beta)$.[25] En adelante —imitando a Gentzen— llamo 1 al ordinal gentzeniano $\omega^0$ y $\omega$ a $\omega^1$.

Gentzen asigna a cada derivación un ordinal de los suyos mediante el procedimiento que paso a explicar. Sea **Δ** una derivación cualquiera. **Δ** consta de uno o más secuentes y rayas horizontales. Asignamos un ordinal a cada secuente y a cada raya, según la posición que ocupa en **Δ** (dos copias del mismo mismo secuente pueden tener distintos ordinales). Llamo $\mathbb{O}(\sigma)$ al ordinal asignado al secuente $\sigma$ (en una determinada posición). Si $\sigma$ es un secuente inicial, $\mathbb{O}(\sigma) = 1$. El ordinal $\mathbb{O}(\lambda)$ de una raya $\lambda$ se determina así: (i) si $\lambda$ pertenece a un debilitamiento, permutación o contracción cuya premisa es un secuente $\sigma$, $\mathbb{O}(\lambda) = \mathbb{O}(\sigma)$; (ii) si $\lambda$ pertenece a un corte con premisas $\sigma_1$ y $\sigma_2$, $\mathbb{O}(\lambda) = \mathbb{O}(\sigma_1) \# \mathbb{O}(\sigma_2)$; (iii) si $\lambda$ pertenece a una inferencia con

[24] Me confieso incapaz de entender cómo un concepto que para nosotros no está definido de ninguna manera (*gar nicht definiert*) puede contribuir a la clarificación (*zur Erläuterung dienen*).

[25] Como el lector habrá advertido, la suma natural así definida se aplica sólo a los ordinales mayores que 0. Pero no cuesta nada incluir al 0 en el dominio de la operación estipulando que $\alpha + 0 = 0$, cualquiera que se $\alpha$. Entonces, la condición $\alpha \# \beta > \alpha$ se cumple sólo si $\beta \neq 0$.

conectivo con una sola premisa $\sigma$, $\mathbb{O}(\lambda) = \mathbb{O}(\sigma)$; (iv) si $\lambda$ pertenece a una inferencia con conectivo con dos premisas $\sigma_1$ y $\sigma_2$, $\mathbb{O}(\lambda) = \max(\mathbb{O}(\sigma_1), \mathbb{O}(\sigma_2))$; (v) si $\lambda$ pertenece a una inferencia inductiva cuya premisa tiene el ordinal $\alpha = \omega^{\alpha_1} + \omega^{\alpha_2} + \ldots + \omega^{\alpha_p}$, $\mathbb{O}(\lambda) = \omega^{\alpha_1+1}$ (naturalmente, si $\alpha_1 = 0$, $\mathbb{O}(\lambda) = \omega^1 = \omega$). Sea $\varsigma$ la conclusión bajo la raya $\lambda$. El ordinal $\mathbb{O}(\varsigma)$ se determina comparando la *altura* $h$ de $\varsigma$ con la altura $h^*$ de las premisas sobre $\lambda$. Si $h = h^*$, $\mathbb{O}(\varsigma) = \mathbb{O}(\lambda)$. Si $h = h^* - 1$, $\mathbb{O}(\varsigma) = \omega^{\mathbb{O}(\lambda)}$. Si $h = h^* - 2$, $\mathbb{O}(\varsigma) = \omega^{\omega^{\mathbb{O}(\lambda)}}$. Si $h = h^* - 3$, $\mathbb{O}(\varsigma) = \omega^{\omega^{\omega^{\mathbb{O}(\lambda)}}}$, etc. La utilidad de estas reglas, al parecer caprichosas, se verá en la próxima etapa del razonamiento. Por ahora, basta que el lector se convenza, analizando ejemplos, de que ellas asignan unívocamente un genuino ordinal a cada secuente de la derivación $\mathbf{\Delta}$. El ordinal $\mathbb{O}(\mathbf{\Delta})$ asignado a la derivación misma es el ordinal de su secuente final.

En la Sección 2.12.2 se demostró que, si existe una derivación fatal $\mathbf{\Delta}$, también existe una derivación fatal $\mathbf{\Delta}^\dagger$, construida a partir de $\mathbf{\Delta}$ mediante la serie de transformaciones que allí se explica. Ahora demostraré que $\mathbb{O}(\mathbf{\Delta}^\dagger) < \mathbb{O}(\mathbf{\Delta})$, examinando dichas transformaciones una a una. Al leer los próximos párrafos conviene tener presentes los párrafos de igual título en la Sección 2.12.2.

*Eliminación de variables libres ociosas*. Dimos por supuesto que $\mathbf{\Delta}$ no contenía ninguna variable libre que no fuera la variable propia de una inferencia, ni dos variables iguales que fueran variables propias de dos inferencias distintas. Esta suposición no afecta el ordinal $\mathbb{O}(\mathbf{\Delta})$. En efecto, si $\mathbf{\Delta}$ no cumple esta doble condición y $\mathbf{\Delta}^{\maltese}$ es la derivación fatal obtenida mediante las sustituciones de variables descritas bajo este mismo título en la Sección 2.12.2, es claro que $\mathbb{O}(\mathbf{\Delta}) = \mathbb{O}(\mathbf{\Delta}^{\maltese})$.

*Eliminación de las inferencias inductivas en el trozo final*. Supongamos que $\mathbf{\Delta}$ contiene inferencias inductivas en su trozo final y que $\mathbf{\Delta}_1$ es la derivación obtenida al eliminar la última de esas inferencias inductivas del modo descrito bajo este mismo título en la Sección 2.12.2. Supongamos que, igual que allí,

$$\frac{\mathfrak{F}(\mathfrak{a}/\mathfrak{x}),\ \Gamma \rightarrow \Theta,\ \mathfrak{F}(\mathfrak{a}'/\mathfrak{x})}{\mathfrak{F}(1/\mathfrak{x}),\ \Gamma \rightarrow \Theta,\ \mathfrak{F}(\mathfrak{n}/\mathfrak{x})}$$

es la forma de esa inferencia (con término numérico $\mathfrak{n} \neq 1$). Sea $\omega^{\alpha_1} + \omega^{\alpha_2} + \ldots + \omega^{\alpha_n}$ el ordinal de la premisa. Entonces, el ordinal de la raya horizontal es $\omega^{\alpha_1+1}$. Ahora bien, la conclusión tiene la misma altura que la premisa, ya que los cortes propios de las alianzas de $\mathfrak{F}(1/\mathfrak{x})$ y $\mathfrak{F}(\mathfrak{n}/\mathfrak{x})$ figuran bajo ambas y tienen por lo menos el mismo grado que la inferencia inductiva considerada. Por lo tanto, $\omega^{\alpha_1+1}$ es también el ordinal de la conclusión. En la derivación $\mathbf{\Delta}_1$, los "secuentes iniciales" del segmento de derivación que sustituye a la inferencia inductiva eliminada se obtienen reemplazando la variable libre $\mathfrak{a}$ por términos numéricos. Como tales reemplazos no afectan el ordinal, esos "secuentes iniciales" tienen todos el mismo ordinal $\omega^{\alpha_1} + \omega^{\alpha_2} + \ldots + \omega^{\alpha_n}$. Los nuevos cortes que aparecen en $\mathbf{\Delta}_1$ tienen todos el mismo grado que la inferencia inductiva eliminada. Por lo tanto, el ordinal del secuente $\mathfrak{F}(1/\mathfrak{x}), \Gamma \rightarrow \Theta, \mathfrak{F}(\mathfrak{n}/\mathfrak{x})$ al final del segmento sustituto es la suma natural de los ordinales de dichos "secuentes iniciales" y su primer término es $\omega^{\alpha_1}$, de tal modo que ese ordinal es menor que $\omega^{\alpha_1+1}$, el ordinal de la conclusión de la inferencia inductiva eliminada. Como ahora se verá, esto implica que $\mathbb{O}(\mathbf{\Delta}_1) < \mathbb{O}(\mathbf{\Delta})$. En efecto, encontraremos bajo el referido secuente $\mathfrak{F}(1/\mathfrak{x}), \Gamma \rightarrow \Theta, \mathfrak{F}(\mathfrak{n}/\mathfrak{x})$ solamente inferencias estructurales que la transformación de $\mathbf{\Delta}$ en $\mathbf{\Delta}_1$ no altera en nada. Los debilitamientos, permutaciones y contracciones trasmiten a la conclusión el ordinal de la premisa. Por lo tanto, si no hay cortes bajo dicho secuente, es claro que $\mathbb{O}(\mathbf{\Delta}_1) < \mathbb{O}(\mathbf{\Delta})$. Por otra parte, un corte cuyas dos premisas tienen ordinales $\alpha$ y $\beta$ trasmite a la conclusión el ordinal $\alpha \# \beta$. Si $\alpha_1 < \alpha$, $(\alpha_1 \# \beta) < (\alpha \# \beta)$. Por lo tanto, aunque haya cortes bajo el referido secuente, $\mathbb{O}(\mathbf{\Delta}_1) < \mathbb{O}(\mathbf{\Delta})$. En la Sección 2.12.2 llamé $\mathbf{\Delta}^0$ a la derivación fatal sin inferencias inductivas en su trozo final obtenida eliminando una a una las inferencias de esa clase en el trozo final de una derivación fatal cualquiera $\mathbf{\Delta}$. Es posible que $\mathbf{\Delta}$ misma no contenga tales inferencias —en cuyo caso $\mathbf{\Delta} = \mathbf{\Delta}^0$— pero, si las contiene, su ordinal disminuye con la eliminación de cada una. Por lo tanto, $\mathbb{O}(\mathbf{\Delta}^0) \leq \mathbb{O}(\mathbf{\Delta})$.

*Eliminación de los debilitamientos y los secuentes básicos lógicos en el trozo final*. Nos toca ahora mostrar que $\mathbb{O}(\mathbf{\Delta}^\star) \leq \mathbb{O}(\mathbf{\Delta}^0)$, donde $\mathbf{\Delta}^\star$ es la derivación fatal sin debilitamientos o secuentes básicos lógicos obtenida por transformación de $\mathbf{\Delta}^0$. Esta fase del razonamiento, aunque elemental, es engorrosa y Gentzen sugiere omitirla si uno se interesa sólo en "lo más esencial" (*das Wesentlichste*—1938, p. 41). Pero en una demostración cualquier fase

es igualmente esencial, a menos que sea superflua (en cuyo caso, se la pue-de suprimir del todo). La transformación de $\mathbf{\Delta}^0$ en $\mathbf{\Delta}^\star$ consta de cero o más etapas, en cada una de las cuales se elimina un debilitamiento o un SBL. Una etapa dada puede envolver la eliminación de un corte o afectar únicamente inferencias estructurales que no sean cortes. En este último caso, el ordinal de la derivación transformada es igual al de la derivación obtenida. Consideremos, pues, sólo el caso especial en que la transformación de $\mathbf{\Delta}^0$ en $\mathbf{\Delta}^\star$ consta de una sola etapa que elimina un corte en el trozo final. Entonces puede ocurrir que disminuya la *altura* de los secuentes situados sobre ese corte, no sólo en el trozo final, sino en la derivación entera.[26] ¿Qué efecto tiene tal reducción de alturas sobre el ordinal de la derivación? Para apreciarlo mejor, y siguiendo el ejemplo de Gentzen, encaramos el asunto así: sea $\sigma$ la premisa del corte eliminado que sobrevive a la transformación; supongamos que la altura de $\sigma$ cae de $h$ a $h$–1 y que esta reducción se propaga, inferencia por inferencia, de conclusiones a premisas, a lo largo de todas las ramas de la derivación que quedan sobre $\sigma$;[27] si la altura de $\sigma$ cae de $h$ a $h$–$n$, entendemos que el proceso se repite $n$ veces; en todo caso, el efecto global será el resultado de los efectos locales de la propagación (repetida o no). Supongamos, entonces, que el "contagio" ha llegado a la conclusión $\varsigma$ de cierta inferencia y pasa ahora a sus premisas $\pi_1$ y $\pi_2$ (a $\pi_1$, si tiene sólo una). Sean $\alpha$ y $\beta$, respectivamente, los ordinales de $\pi_1$ y $\pi_2$ antes de que esto ocurra. Si $\pi_1$ y $\pi_2$ son secuentes iniciales, $\alpha = \beta = 1$ y no cambian con la reducción de altura. Por lo tanto, sólo nos interesa el caso en que $\pi_1$ o $\pi_2$ o ambas son conclusiones de inferencias. En tal caso, cuando el "contagio" las alcanza, sus ordinales se transforman de $\alpha$ en $\omega^\alpha$ y de $\beta$ en $\omega^\beta$. El ordinal de la raya horizontal, que, según el tipo de la inferencia en cuestión, era $\alpha$, o $\alpha \# \beta$, o $\max(\alpha+1,\beta+1)$, o $\omega^{\alpha_1+1}$ (si se trata de una inferencia inductiva y $\alpha = \omega^{\alpha_1} + \omega^{\alpha_1} + \ldots + \omega^{\alpha_n}$), se convierte, respectivamente, en $\omega^\alpha$, o en $\omega^\alpha \# \omega^\beta$, o en $\max(\omega^\alpha+1,\omega^\beta+1)$, o sigue siendo igual a $\omega^{\alpha_1+1}$. ¿Qué pasa con el ordinal de la conclusión $\varsigma$? Si antes del "contagio" la diferencia de altura entre $\varsigma$ y sus premisas era 1, y se ha reducido, por ende, a 0, el ordinal de

[26] La reducción de altura ocurre si y sólo si el corte eliminado tenía un grado $g$ mayor que el de cualquier otro corte situado bajo él.

[27] La propagación hacia arriba llega hasta la conclusión del primer corte o inferencia inductiva con grado $g \geq h$, pero no pasa a sus premisas, cuya altura, por cierto, sigue siendo $g$.

$\varsigma$ después del "contagio" sigue siendo $\omega^{\alpha}$, o se convierte de $\omega^{\alpha \# \beta}$ en $\omega^{\alpha} \# \omega^{\beta}$, o de $\max(\omega^{\alpha+1},\omega^{\beta+1})$ en $\max(\omega^{\alpha}+1,\omega^{\beta}+1)$, o de $\omega^{\omega^{\alpha+1}}$ en $\omega^{\omega^{\alpha_1}+\ldots+1}$. En cada caso, pues, el ordinal de $\varsigma$ es igual o menor que lo que era antes. Otro tanto cabe decir si la diferencia entre $\varsigma$ y sus premisas era $n > 1$ y pasa a ser $n$–1 (verifíquelo el lector). Este análisis demuestra que, cuando ocurre reducción de alturas en el curso de la transformación de $\mathbf{\Delta}^0$ en $\mathbf{\Delta}^{\star}$, nunca aumenta el ordinal de un secuente situado bajo otro cuya altura disminuye. En particular, no puede aumentar el ordinal del secuente final, que es el ordinal de la derivación. Si la transformación de $\mathbf{\Delta}^0$ en $\mathbf{\Delta}^{\star}$ tiene más de una etapa, este resultado se aplica a cada una. Por lo tanto, $\mathbb{O}(\mathbf{\Delta}^{\star}) \leq \mathbb{O}(\mathbf{\Delta}^0)$.

*La reducción del conectivo*. Hemos logrado establecer que, si $\mathbf{\Delta}$ es una derivación fatal, hay derivaciones fatales $\mathbf{\Delta}^0$ y $\mathbf{\Delta}^{\star}$, con las propiedades señaladas, cuyo ordinal es *igual o menor* que el ordinal de $\mathbf{\Delta}$. Mas para demostrar inductivamente que no hay ninguna derivación fatal es preciso comprobar que, si hubiera una, existiría también otra cuyo ordinal es *estrictamente menor* que el de la primera. La artificiosa operación que llamamos "reducción del conectivo" garantiza justamente este resultado. Como sabemos, esta operación tiene que ser aplicable a una derivación fatal con las propiedades de $\mathbf{\Delta}^{\star}$, si tal derivación existe, porque una derivación así tendría que incluir *por lo menos una inferencia con conectivo*. Atendamos, pues, a la reducción del conectivo aplicada al signo de cuantificación $\forall$, ilustrada en los Cuadros 1 y 2 de la Sección 2.12.2. Para facilitar las referencias, marqué allí ciertas líneas con letras acompañadas del número del cuadro respectivo. Llamamos $\mathbf{\Delta}^{\star}$ y $\mathbf{\Delta}^{\dagger}$ las derivaciones representadas en los Cuadros 1 y 2, respectivamente. La líneas $[c_1]$ y $[d_2]$ son las primeras en que aparecen secuentes cuya altura es menor que la altura de las premisas de los cortes en $[b_1]$ y $[c_2]$. Sea $\alpha$ el ordinal de la raya horizontal sobre $[c_1]$ y sean $\beta$ y $\gamma$ los ordinales de las rayas horizontales sobre $[d_2]$. Suponemos que $\beta \geq \gamma$ ($\beta$ puede corresponder a la raya de la derecha). Entonces $\alpha > \beta \geq \gamma$. Esta desigualdad se basa en lo siguiente: (i) las alturas de los secuentes que preceden inmediatamente a dichas rayas horizontales tienen el mismo valor —llamémoslo $\rho$— en $\mathbf{\Delta}^{\star}$ y $\mathbf{\Delta}^{\dagger}$;[28] (ii) sobre la raya en $\mathbf{\Delta}^{\star}$ hay una inferencia con conectivo *más* que sobre

[28] En efecto, dicha altura es el grado más alto de un corte bajo esas rayas (como se recordará, no hay ninguna inferencia inductiva bajo ellas). Ahora bien, salvo por el nuevo corte con grado 0 entre $[d_2]$ y $[e_2]$, todos los cortes bajo $[c_2]$ reproducen cortes bajo $[b_1]$.

cada una de las rayas en $\mathbf{\Delta}^\dagger$; (iii) si el ordinal de la premisa de esa inferencia con conectivo es $\mu$, el de su raya horizontal es $\mu$+1 y ese 1 adicional se trasmite a través de las inferencias —todas estructurales— que llevan de las conclusiones anotadas en [$a_1$] a las anotadas en [$b_1$]; (iv) dicha inferencia con conectivo ha sido reemplazada en $\mathbf{\Delta}^\dagger$ con inferencias estructurales cuyas rayas horizontales tienen el mismo ordinal que las respectivas premisas. Como ahora veremos, la desigualdad $\alpha > \beta \geq \gamma$ implica que el secuente $\Gamma_3 \rightarrow \Theta_3$ en la línea [$g_2$] tiene un ordinal menor que el secuente $\Gamma_3 \rightarrow \Theta_3$ en la línea [$c_1$]. Este último es el primer secuente bajo el corte propio de la alianza de la fórmula $\forall\mathfrak{x}\mathfrak{F}$ cuya altura —llamémosla *u*— es menor que la altura *r* de las premisas de ese corte. Como $\mathbf{\Delta}^\star$ y $\mathbf{\Delta}^\dagger$ son iguales bajo dicho secuente $\Gamma_3 \rightarrow \Theta_3$ es claro que la altura de éste en [$g_2$] también es *u*. Asimismo, *u* es la altura del secuente $\Gamma_3, \Gamma_3 \rightarrow \Theta_3, \Theta_3$ en la línea [$f_2$], puesto que entre [$f_2$] y [$g_2$] no hay ningún corte o inferencia inductiva. Entre las líneas [$e_2$] y [$f_2$] ocurre el *nuevo corte*. Sea *t* la altura de sus premisas. Es claro que $u \leq t$. Comprobemos que $t < r$. En efecto, o bien $t = u < r$, o bien *t* es igual al grado de la formula $\mathfrak{F}(\mathfrak{n}/\mathfrak{x})$ cortada en el nuevo corte; ahora bien, *r* es mayor o igual que el grado de $\forall\mathfrak{x}\mathfrak{F}$, el cual, por cierto, es mayor que el grado de $\mathfrak{F}(\mathfrak{n}/\mathfrak{x})$. Supongamos por el momento que las diferencias entre las alturas *r*, *t* y *u* son mínimas, esto es, que $r = t + 1$ y $t = u$. Como el ordinal de la raya horizontal sobre la línea [$c_1$] es $\alpha$, el ordinal de $\Gamma_3 \rightarrow \Theta_3$ en $\mathbf{\Delta}^\star$ es $\omega^\alpha$. Como los ordinales de las rayas horizontales sobre [$d_2$] son $\beta$ y $\gamma$ ($\beta \geq \gamma$), los ordinales de los secuentes en [$d_2$] son $\omega^\beta$ y $\omega^\gamma$. Estos ordinales se trasmiten inalterados a través de las permutaciones entre [$d_2$] y [$e_2$]. Por lo tanto, una de las premisas del nuevo corte tiene el ordinal $\omega^\beta$ y la otra el ordinal $\omega^\gamma$, de modo que el ordinal de la conclusión del nuevo corte es $\omega^\beta \# \omega^\gamma$, igual al ordinal de $\Gamma_3 \rightarrow \Theta_3$ en $\mathbf{\Delta}^\dagger$. Ahora bien, $\alpha > \beta \geq \gamma$ implica que $\omega^\alpha > \omega^\beta + \omega^\gamma = \omega^\beta \# \omega^\gamma$.[29] Si las diferencias entre las alturas *r*, *t* y *u* exceden

29 Sea $\mathfrak{S}_m$ el "sistema" de índice más bajo a que pertenecen $\alpha$, $\beta$ y $\gamma$. Utilizando la definición de la relación < en $\mathfrak{S}_{m+1}$ es claro que $\omega^\alpha > \omega^\beta + \omega^\gamma$, puesto que el primer (y único) "exponente" del ordinal de la izquierda es mayor que el primer "exponente" del ordinal de la derecha. La desigualdad también puede probarse apelando al isomorfismo entre los ordinales de Gentzen y un segmento de los ordinales cantorianos, puesto que, si $\alpha$, $\beta$, $\gamma$ y $\nu$ son ordinales cantorianos tales que $\alpha > \beta \geq \gamma$ y $\nu \geq 3$, $\nu^\alpha \geq \nu\cdot\nu^\beta > 2\nu^\beta \geq \nu^\beta + \nu^\gamma$. Pero esta prueba utiliza las propiedades de la multiplicación y la exponienciación de ordinales, que sólo pueden definirse por inducción transfinita (cf. Apéndice VIII).

el mínimo, esto es, si $r > t + 1$ o $t > u$, en vez de la desigualdad $\omega^\alpha > \omega^\beta + \omega^\gamma$ se usa la desigualdad

$$\omega^{\cdot^{\cdot^{\omega^{\cdot^{\cdot^{\cdot^{\omega^\alpha}}}}}}} > \omega^{\cdot^{\cdot^{\omega^{\cdot^{\cdot^{\cdot^{\omega^\beta}}}}}}} + \omega^{\cdot^{\cdot^{\cdot^{\omega^\gamma}}}}$$

Como $\mathbf{\Delta}^\star$ y $\mathbf{\Delta}^\dagger$ son idénticas bajo el secuente $\Gamma_3 \rightarrow \Theta_3$, la desigualdad entre los ordinales de este secuente se trasmite al secuente final. Por lo tanto $\mathbb{O}(\mathbf{\Delta}^\dagger) < \mathbb{O}(\mathbf{\Delta}^\star)$.

Hemos demostrado, pues, que si $\mathbf{\Delta}$ es una derivación fatal, hay derivaciones fatales, $\mathbf{\Delta}^0$, $\mathbf{\Delta}^\star$ y $\mathbf{\Delta}^\dagger$ tales que $\mathbb{O}(\mathbf{\Delta}^\dagger) < \mathbb{O}(\mathbf{\Delta}^\star) \leq \mathbb{O}(\mathbf{\Delta}^0) \leq \mathbb{O}(\mathbf{\Delta})$. Este resultado lleva derechamente a la conclusión de que no hay derivaciones fatales. Razonamos inductivamente, como sigue.

(I) Si $\mathbf{\Delta}$ es una derivación tal que $\mathbb{O}(\mathbf{\Delta}) = 1$, $\mathbf{\Delta}$ no incluye cortes, inferencias inductivas ni inferencias con conectivo.[30] Consta, pues, de un solo hilo y su secuente final reproduce las prefórmulas y posfórmulas del secuente inicial (posiblemente acompañadas de otras, añadidas por debilitamiento). Por lo tanto, $\mathbf{\Delta}$ no es una derivación fatal.

(II) Sea $\alpha$ un ordinal de Gentzen mayor que 1. La hipótesis inductiva dice que si $\mathbf{\Gamma}$ es una derivación tal que $\mathbb{O}(\mathbf{\Gamma}) < \alpha$, $\mathbf{\Gamma}$ no es una derivación fatal. Sea $\mathbf{\Delta}$ es una derivación tal que $\mathbb{O}(\mathbf{\Delta}) = \alpha$. Entonces $\mathbf{\Delta}$ no puede ser fatal. En efecto, si lo fuera, existiría una derivación fatal $\mathbf{\Delta}^\dagger$ tal que $\mathbb{O}(\mathbf{\Delta}^\dagger) < \mathbb{O}(\mathbf{\Delta}) = \alpha$, contradiciendo la hipótesis inductiva.

(III) Como toda derivación tiene un ordinal de Gentzen igual o mayor que 1, (I) y (II) implican que ninguna derivación es fatal.

En virtud del isomorfismo entre los ordinales de Gentzen y los ordinales de Cantor, $\alpha$ recorre toda la serie de los ordinales entre 1 y $\varepsilon_0$. Así, el argumento enterior es una inducción transfinita hasta el ordinal cantoriano $\varepsilon_0$. En

30 Como se recordará, en estas tres clases de inferencia, el ordinal de la conclusión es mayor que el de las premisas y, por lo tanto, es mayor que 1. En cambio, las inferencias por debilitamiento, contracción y permutación trasmiten a la conclusión el ordinal de la premisa.

este respecto, Gentzen subraya que si prescindimos de las inferencias inductivas (de la clase IM; p. 428), la consistencia del cálculo puede establecerse por su método sin recurrir a una inducción transfinita. Para ello, hay que hacer los siguientes cambios en la asignación de ordinales a secuentes y derivaciones: (i) escribir el número 3 donde antes escribíamos ω;[31] (ii) reemplazar "suma natural" por "suma" (la operación aritmética ordinaria); (iii) entender las relaciones < y =, así como las operaciones de multiplicación y exponenciación, en su sentido aritmético ordinario (aplicado a números naturales). Si no hay inferencias IM en el cálculo, el argumento inductivo para probar su consistencia puede utilizar el orden así definido. Entonces, obviamente, el ordinal α mencionado en la hipótesis inductiva (II) es un número natural y la conclusión (III) se obtiene por inducción ordinaria. Vemos así que la inducción transfinita hasta el ordinal cantoriano $\varepsilon_0$ se requiere sólo para probar la consistencia del cálculo *con* inferencias IM. Ello no puede sorprendernos. Al fin y al cabo, la escuela de Hilbert había establecido antes de 1930, con métodos incuestionablemente finitistas, la consistencia de la aritmética *sin* el Axioma P5 de Peano (que valida la inferencia inductiva ordinaria). Por otra parte, en vista de los resultados de Gödel (1931), no es posible probar la consistencia de un sistema de aritmética formalizada con argumentos formalizables dentro de ese sistema, y la inducción matemática ordinaria es obviamente formalizable en el cálculo de Gentzen si se admiten las inferencias conforme al esquema IM. En suma, la prueba de Gentzen excede los recursos corrientes de la metamatemática hilbertiana para lograr un propósito que ya se sabía que era inalcanzable con ellos. A este respecto, vale la pena anotar que Hilbert y Bernays (1939, §§5 y 3c) justificaron *con los recursos de la aritmética finitista* la inducción transfinita hasta cualquier ordinal menor que $\varepsilon_0$,[32] y que Gentzen (1943) estableció que la inducción transfinita hasta $\varepsilon_0$ no puede justificarse con dichos recursos. Esto quiere decir que sus dos demostraciones de la consistencia de la aritmética (Gentzen 1935, 1938) exceden los recursos de la aritmética finitista sólo el mínimo indis-

31 Se usa el 3 por la razón indicada en la nota 29: 3 es el más pequeño número natural $n$ tal que si $a, b, c \in \mathbb{N}$ y $a > b \geq c$, $n^a \geq n \cdot n^b > 2n^b \geq n^b + n^c$. Como vimos, esta desigualdad se invoca en una etapa del razonamiento.

32 Gentzen aparentemente alude a este resultado cuando observa que la inducción transfinita hasta un ordinal menor que $\varepsilon_0$ se puede demostrar en su cálculo (1938, p. 43).

33 El resultado de Gentzen 1943 puede verse como una confirmación independiente de

pensable para no contradecir los resultados de Gödel (1931).[33]

Alfred Tarski observó en una ocasión que, aunque la prueba de consistencia de Gentzen es sin duda muy interesante y puede resultar muy estimulante y fecunda, no le parecía que "la consistencia de la aritmética fuese ahora mucho más evidente para mí que antes de conocer esa prueba" (1954, p. 19; cit. en Szabo 1969, p. 10). Pero Gentzen no pretendía demostrar la consistencia de la *aritmética*. Como él mismo explica con insuperable claridad:

> No se trata en absoluto de la consistencia (*Widerspruchsfreiheit*) del *dominio de los números* como tal, esto es, de las relaciones básicas entre los números fijadas por los "axiomas" (por ejemplo, por los "Axiomas de Peano" de la aritmética). Demostrar la consistencia de estos axiomas sin dar ya por supuesto algo equivalente es impensable. Se trata más bien de la consistencia de la *deducción lógica* (*des logischen Schließens*) aplicada a los números naturales (*a partir* de sus *axiomas*), como se la practica en las *pruebas* de la aritmética.
>
> (Gentzen 1935, p. 498)

Concretamente, se trata de establecer que no es posible derivar contradicciones en la aritmética *formalizada* en el cálculo de Gentzen, o en otro cálculo equivalente o menos poderoso. Para la escuela de Hilbert este resultado era imprescindible, pues sólo sobre esa base se podía establecer la consistencia de la teoría de conjuntos formalizada, recuperando así todas las comodidades que ofrece al matemático "el paraíso de Cantor" sin suscribir la metafísica cantoriana del infinito. En lo relativo al infinito, la postura de Gentzen era bien decidida:

> *No se debe considerar un agregado infinito como cerrado y existente en sí (infinitud actual), sino sólo como algo en devenir, que puede siempre seguir edificándose constructivamente desde lo finito (infinitud potencial).*
>
> (Gentzen 1935, p. 524; cursiva en el original)[34]

---

Gödel 1931, que ilustra "de un modo *directo* la incompletud del formalismo aritmético" (Szabo 1969, p. 17). Después de Gentzen, Ackermann (1940) y Schütte (1951) han publicado nuevas pruebas de la consistencia de la aritmética elemental que, por cierto, también se valen de la inducción transfinita hasta $\varepsilon_0$.

34 Hablando de la tesis intuicionista de que las proposiciones sobre el infinito actual carecen de sentido y deben, por ende, rechazarse aunque sean consistentes, Gentzen declara: "Creo, por ejemplo, que en la teoría general de conjuntos una investigación

Por eso mismo, como luego veremos, pone gran empeño en convencernos de que su prueba de consistencia sólo emplea razonamientos finitistas, conforme a las exigencias de Hilbert. Pero esa prueba debe juzgarse valiosa también aparte de todo melindre metafísico, especialmente ahora que tantos resultados matemáticos son el producto de computaciones electrónicas que sólo pueden verificarse cotejándolas con otras computaciones. Tal cotejo sólo certifica —o, mejor dicho, hace sumamente probable— que las computaciones son *correctas*, que las máquinas que las ejecutaron no han cometido errores en la ejecución de las instrucciones del programa. Pero el resultado de una computación correcta puede ser incorrecto si el programa mismo es la versión electrónica de un cálculo inconsistente. De ahí que sea importante probar la consistencia del cálculo incorporado en el programa, con métodos evidentes para nosotros, aunque no sean formalizables en ese cálculo.

Aunque las alegaciones de Gentzen no acaben de persuadirnos de que su inducción transfinita merece llamarse 'finita' o siquiera 'finitista' (*finit*), explican con gran claridad por qué tiene que resultarnos —y de hecho nos resulta— evidente. Gentzen admite que, a diferencia de todos los otros pasos de su prueba, la inducción transfinita no es propiamente elemental. Por eso "encaramos aquí una *tarea de índole completamente distinta*: lo esencial no es *demostrar* la inducción transfinita —lo cual no es difícil y se puede

---

metamatemática cuidadosa confirmará eventualmente que todas las cardinalidades mayores que las numerables son en un sentido muy definido sólo una apariencia vacua y que sería prudente prescindir de estos conceptos" (1936/37, en Gentzen, CP, p. 229).

35 Gentzen escribe "vielmehr sie auf *finiter* Grundlage zu beweisen". Para el lector alemán desapercibido *finit* no puede significar otra cosa que 'finito, finita'. Pero entonces la palabra tiene un significado inequívoco que no puede cubrir la inducción transfinita de Gentzen, no sólo a causa de su nombre de mala fama ("ihres anrüchigen Namens wegen"—Gentzen 1938, p. 44), sino porque ella recorre *toda* la serie de los números naturales una multitud infinita —aunque numerable— de veces. Pero Gentzen dice expresamente que el carácter 'finit' de su prueba no puede propiamente "demostrarse" debido a que el concepto *finit* no está inequívocamente definido y malamente se puede definir ("weil der Begriff 'finit' nicht eindeutig formal abgegrenzt ist und auch kaum abgegrenzt werden kann"—1935, p. 557). Por eso, prefiero traducir *finit* con 'finitista' (los autores de habla inglesa dicen 'finitary') y dejar el significado de esta palabra en la penumbra que el programa de Hilbert aparentemente necesita para sobrevivir.

hacer de varias maneras— sino más bien demostrarla sobre una base *finitista*,[35] esto es, poner en claro que es un modo de inferencia acorde con el principio de la *concepción constructiva* del infinito" (1938, p. 44). Para que no quepa duda de esto, Gentzen subraya reiteradamente que sus ordinales no son lo mismo que los ordinales de Cantor, sino unos objetos "formales" —esto es, símbolos, figuras que pueden escribirse en un papel o en una pizarra— definidos por él. Gentzen estima que su definición es estrictamente constructiva: cada ordinal es una figura formada según reglas precisas con los dos signos primarios 0 y $\omega$. Dado un ordinal cualquiera $\alpha$, sabemos construir su sucesor inmediato $\alpha + \omega^0$. Pero ¿es justo llamar "constructivo" el paso de cada sistema de ordinales $\mathfrak{S}_n$ al sistema siguiente $\mathfrak{S}_{n+1}$? Dicho paso *presupone* que los ordinales del sistema $\mathfrak{S}_n$ *ya estén definidos*, y el número de éstos es obviamente *infinito* (cf. p. 441, nota 23). Parecería que, contrariando el citado precepto de Gentzen, el sistema infinito $\mathfrak{S}_n$ se tiene que considerar aquí *como cerrado*, y no *sólo como algo en devenir.* Si este modo de introducir sistema $\mathfrak{S}_{n+1}$ (para cada $n \in \mathbb{N}$) se ajusta a la "concepción constructiva del infinito", ¿en qué se aparta de ella la definición cantoriana de los ordinales? Para responder a esta pregunta, conviene recordar los tres "principios generadores" aducidos por Cantor (GA, pp. 195ss; *vide supra*, Capítulo 2.5). El más aventurado de ellos, el principio "de inhibición o limitación", produce los ordinales cantorianos de las clases tercera y superior, y por lo tanto no desempeña ningún papel en la definición de los ordinales de Gentzen. Los otros dos presiden la generación (i) del ordinal siguiente a cada ordinal dado y (ii) del ordinal límite de cada serie numerable infinita de ordinales sucesivos. En virtud del isomorfismo entre los ordinales de Cantor ($< \varepsilon_0$) y los de Gentzen, es claro que éstos satisfacen tanto el principio (i) como el principio (ii). Pero los ordinales de Gentzen no se *generan* según el principio (ii) —que un finitista debiera cuestionar—, sino según una regla que autoriza a *escribir* expresiones de la forma $\omega^{\alpha_1} + \omega^{\alpha_2} + \ldots + \omega^{\alpha_m}$ con cualquier número de exponentes $\alpha_k$ seleccionados arbitrariamente en un sistema de expresiones dado. Una regla como ésta, puramente sintáctica, es sin duda constructiva. Para el finitista no hay pues ninguna dificultad en el *tránsito* de un sistema gentzeniano dado $\mathfrak{S}_n$ al sistema siguiente $\mathfrak{S}_{n+1}$. Pero a la luz de esta explicación resalta aún más la dificultad señalada arriba: si exceptuamos el sistema $\mathfrak{S}_0$ cuyo único elemento es el objeto 0, ¿qué sistemas de Gentzen están *dados*? El Axioma del Infinito de los conjuntistas nos da,

por cierto, el sistema $\mathfrak{S}_1$, y una vez que disponemos de él, no cuesta mucho asegurarse los siguientes. Pero el Axioma del Infinito postula justamente un agregado infinito actual, "cerrado y existente en sí". ¿Puede uno aceptarlo y seguir llamándose *finitista*? Y si uno acepta ese agregado y, como es habitual, lo llama ω (en vez de concebir a este signo "de un modo enteramente formal", como pide Gentzen), ¿por qué inhibirse de aceptar $\mathcal{P}\omega$, $\mathcal{P}^2\omega$, . . . , $\mathcal{P}^\omega\omega$, . . . ? ¿Dónde se tira la raya?[36]

Después de Gentzen, la teoría de la prueba se consolida como una disciplina matemática que produce resultados interesantes mediante el análisis combinatorio de la deducibilidad en los cálculos lógicos y las teorías matemáticas formalizadas en ellos (cf. los manuales de Schütte 1960 y Takeuti 1975); pero ya nadie la presenta a la manera de Hilbert como una fuente de certeza filosóficamente incontrovertible, garantía de las ciencias y pilar de la civilización (cf. p. 120, nota 10). En un artículo de enciclopedia publicado en 1980,[37] Schütte describe la teoría de la prueba (*Beweistheorie*) como *una* de las divisiones de la "metamatemática" o investigación matemáticamente exacta de teorías matemáticas formalizadas (se recordará que Hilbert usaba *Beweistheorie* y *Metamathematik* como sinónimos). Schütte asigna a otra división las cuestiones de computabilidad y decidibilidad que hemos visto surgir de la problemática hilbertiana pero que ahora se estudian independientemente, en la teoría de las funciones recursivas y de los conjuntos recursivamente enumerables (cf. los manuales de Rogers 1967 y Soare 1987), orientada sobre todo hacia las "ciencias de la computación". Por último, Schütte incluye en la metamatemática una tercera división: "la teoría de modelos, que se ocupa con las interpretaciones y las cuestiones de realizabilidad de los sistemas de axiomas de la lógica y la matemática". Como veremos en la Parte 3, el enfoque y los métodos de la teoría de modelos han producido desde los años 30 algunos hallazgos sumamente significativos para la filosofía matemática conjuntista. Veremos allí que esta forma de "metamatemática", tal como la cultivan Tarski y Gödel, descarta el prejuicio finitista de Hibert y aprovecha sin remilgos las facilidades de la teoría de conjuntos. En esta

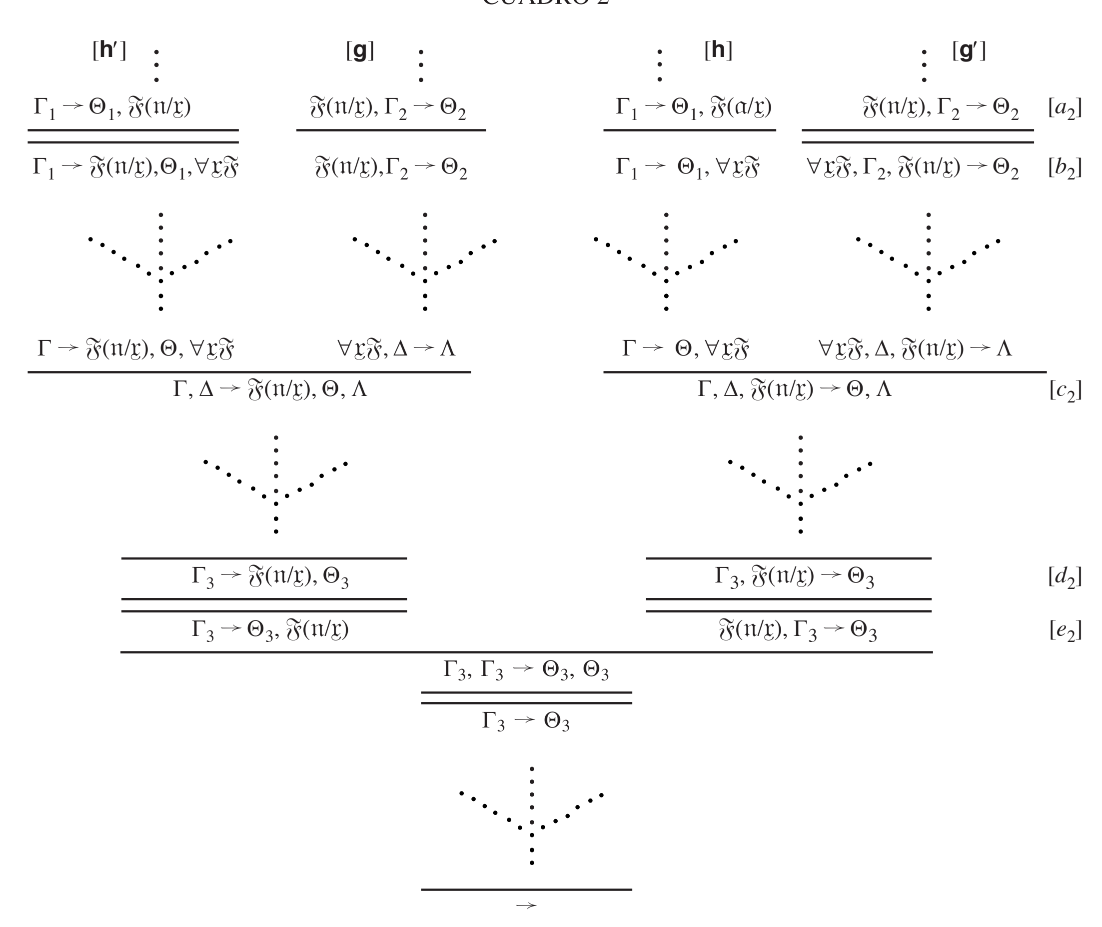
36 Recomiendo leer las iluminadoras observaciones de Gödel sobre la inducción hasta $\varepsilon_0$ (1958, pp. 281 s.; traducción inglesa en Gödel, CW, II, 243).

37 Ritter y Gründer, HWP, vol. 5, col. 1176, *s.v.* 'Metamathematik, II'.

forma, la “metamatemática” explora y perfecciona el paraíso de Cantor desde dentro, en vez de pretender cimentarlo y fortificarlo con recursos pedestres desde fuera. La empresa alcanza su punto más alto cuando Gödel (1938, 1939; cf. 1940) demuestra que la teoría de conjuntos de Zermelo-Fraenkel (ZF) es compatible con el Axioma de Selección (**AS**) y la Hipótesis generalizada del continuo (**HGC**). Quedó así establecido que las conjeturas más audaces de Cantor —el teorema del buen orden, la hipótesis del continuo— son al menos tan viables como el sistema ZF, cuya consistencia se acepta, en general, sin discusión.[38] El conjuntismo triunfante tendrá, sin embargo, un vuelco sorprendente cuando Paul J. Cohen (1963/64, cf. 1966) demuestre, mediante un novedoso método de construcción de modelos, que ZF es compatible también con la negación de **AS** y la negación de **HGC**. Aunque aceptemos que la citada demostración de Gödel prácticamente inmuniza al cantorismo contra el riesgo de contradicción, ¿de qué vale esta seguridad, si el resultado obtenido por Cohen muestra que el de Cantor no es el único paraíso conjuntista disponible y no tenemos cómo saber en cuál estamos parados?

38 Como ha dicho Boolos, “es apenas concebible que ZF sea inconsistente” (1987, en Demopoulos, 1995, p. 219).

CUADRO 2



[h′]
[g]
[h]
[g′]
Γ1 → Θ1, 𝔉(𝔫/𝔵)
𝔉(𝔫/𝔵), Γ2 → Θ2
Γ1 → Θ1, 𝔉(𝔞/𝔵)
𝔉(𝔫/𝔵), Γ2 → Θ2
[a2]
Γ1 → 𝔉(𝔫/𝔵),Θ1,∀𝔵𝔉
𝔉(𝔫/𝔵),Γ2 → Θ2
Γ1 → Θ1, ∀𝔵𝔉
∀𝔵𝔉, Γ2, 𝔉(𝔫/𝔵) → Θ2
[b2]
Γ → 𝔉(𝔫/𝔵), Θ, ∀𝔵𝔉
∀𝔵𝔉, Δ → Λ
Γ → Θ, ∀𝔵𝔉
∀𝔵𝔉, Δ, 𝔉(𝔫/𝔵) → Λ
Γ, Δ → 𝔉(𝔫/𝔵), Θ, Λ
Γ, Δ, 𝔉(𝔫/𝔵) → Θ, Λ
[c2]
Γ3 → 𝔉(𝔫/𝔵), Θ3
Γ3, 𝔉(𝔫/𝔵) → Θ3
[d2]
Γ3 → Θ3, 𝔉(𝔫/𝔵)
𝔉(𝔫/𝔵), Γ3 → Θ3
[e2]
Γ3, Γ3 → Θ3, Θ3
Γ3 → Θ3
→



# APÉNDICES

# APÉNDICES

## I

## LAS DEFINICIONES CANTORIANAS DE ‘CONJUNTO BIEN ORDENADO’

Demostraré aquí que las dos definiciones de conjunto bien ordenado que dio Cantor son equivalentes (v. Capítulo 1.5). Recordemos que un conjunto $M$ es un conjunto linealmente ordenado si y sólo si hay una relación binaria $<$ definida sobre $M$, tal que si $a$, $b$ y $c$ son distintos elementos de $M$ (i) o bien $a < b$ o bien $b < a$, pero nunca las dos cosas a la vez; (ii) si $a < b$ y $b < c$, $a < c$; y (iii) nunca ocurre que $a < a$.

Formularé las condiciones características de un conjunto bien ordenado según ambas definiciones en un cálculo predicativo de primer orden con identidad =, predicados binarios $\in$, $\subseteq$, $<$, y una constante individual $M$ que designa el conjunto de referencia.[1] Para simplificar la escritura empleo cuantificadores restringidos. Esto quiere decir que la expresión ‘$(\forall x\subseteq M)\Phi(x)$’ debe leerse ‘todo objeto $x$ que sea parte de $M$ satisface la condición $\Phi$’ y la expresión ‘$(\exists y\in x)\Phi(y)$’ debe leerse ‘hay un elemento $y$ del conjunto $x$, tal que $\Phi(y)$’. (En suma, $(\forall x\subseteq M)\Phi(x)$ y $(\exists y\in x)\Phi(y)$ abrevian, respectivamente, a $\forall x(x \subseteq M \rightarrow \Phi(x))$ y $\exists y(y \in x \wedge \Phi(y)))$.

Buen Orden (Cantor 1899). Un conjunto linealmente ordenado $M$ está bien ordenado por la relación $<$ si y sólo si toda parte no vacía de $M$ tiene un primer elemento (respecto al orden inducido por $<$), esto es, si y sólo si $M$ cumple la siguiente condición:

BO $\quad (\forall x\subseteq M)(x \neq \varnothing \rightarrow (\exists y\in x)(\forall z\in x)(z \neq y \rightarrow y < z))$

[1] En el Apéndice IX se explica el concepto de cálculo predicativo de primer orden para beneficio de los lectores con cierta educación matemática que no hayan estudiado lógica.

BUEN ORDEN (CANTOR 1883). Un conjunto linealmente ordenado $M$ está bien ordenado por la relación $<$ si y sólo si $M$ cumple las tres condiciones siguientes:

BO1 $(\exists x \in M)(\forall y \in M)(x \neq y \rightarrow x < y)$

(*M* tiene un primer elemento con respecto al orden inducido por <).

BO2 $(\forall x \in M)((\forall y \in M)(y \neq x \rightarrow y < x)) \vee$
$(\exists y \in M)(x < y \wedge (\forall z \in M)(x < z \wedge z \neq y \rightarrow y < z)))$

(Todo elemento de *M* que no sea el último elemento de *M* tiene un sucesor inmediato en el orden inducido por <).

BO3 $(\forall u \subseteq M)(u \neq \varnothing \rightarrow (\exists y \in M)((\forall z \in M)(z \in u \rightarrow z < y) \wedge$
$(\forall w \in M)((z \in u \rightarrow z < w) \rightarrow (y = w \vee y < w)))$

(Si una parte no vacía *u* de *M* tiene sucesores en el orden inducido por < —esto es, si hay elementos de *M* que suceden a todos los elementos de *u*—, *u* tiene un sucesor próximo — esto es, hay un elemento de *M* precedido por todos los elementos de *u* pero a su vez precede a cualquier otro elemento de *M* precedido por todos los elementos de *u*).

Paso a demostrar que si $M \neq \varnothing$, la condición BO equivale a la conjunción de BO1, BO2 y BO3.

BO implica BO1, puesto que $M$ es una parte no vacía de $M$ y por ende tiene un primer elemento. Además, BO implica BO3. En efecto, sea $u$ una parte propia no vacía de $M$ y sea $u'$ su complemento en $M$. Sea $S(u) = \{w{:}z \in u \rightarrow z < w\}$. Entonces, si $S(u) \neq \varnothing$, la intersección $u' \cap S(u) \neq \varnothing$ y tiene un primer elemento $y$, que es el sucesor próximo de $u$. Si $x$ es un elemento de $M$ que no es el último elemento de $M$, el mismo argumento prueba que el conjunto $\{x\}$ tiene un sucesor próximo, que es, por cierto, el sucesor inmediato de $x$. Por lo tanto, BO implica BO2.

Por otra parte, si $M$ no cumple la condición BO hay por lo menos una parte no vacía que no tiene un primer elemento, esto es, un conjunto $u \subseteq M$ tal que $u \neq \varnothing$ y para todo $x \in u$ hay un $y \in u$ tal que $y < x$. Sea $[u]$ el conjunto de los predecesores de $u$, es decir, sea $[u] = \{z \in M{:}\ (\forall x \in M)(x \in u \rightarrow z < x)\}$. Si $[u] = \varnothing$, $M$ no tiene un primer elemento y, por ende, no cumple la condición BO1. Si $[u] \neq \varnothing$, $[u]$ es una parte no vacía de $M$ que

tiene sucesores pero no tiene un sucesor próximo, de modo que $M$ no cumple la condición BO3.

## II

## MÁS SOBRE EL BUEN ORDEN

Las definiciones y teoremas contenidos en este apéndice completan y precisan las ideas sobre conjuntos bien ordenados expuestas en el Capitulo 1.5.

DEFINICIÓN 1. Si los conjuntos $K_1$ y $K_2$ están ordenados, respectivamente, por las relaciones $<_1$ y $<_2$, diremos que la aplicación $f\colon K_1 \to K_2$ es un *morfismo de órdenes* si para todo $x, y \in K_1$, $x <_1 y$ si y sólo si $f(x) <_2 f(y)$. Si $f$ además es biyectiva decimos que es un *isomorfismo de órdenes*. Un isomorfismo cuyo dominio y codominio coinciden se llama un *automorfismo*. (Cuando ello no dé lugar a confusión, omitiré la frase 'de órdenes' y diré sencillamente 'morfismo', 'isomorfismo', 'automorfismo'.)

DEFINICIÓN 2. Si $\langle K,<\rangle$ es un conjunto bien ordenado y $x \in K$, la *sección* determinada por $x$ en $K$ es el conjunto bien ordenado $\langle\{z\colon z < x\},<\rangle$ y el *remanente* determinado por $x$ en $K$ es el conjunto bien ordenado $\langle\{z\colon x \leq z\},<\rangle$.

PROPOSICIÓN 1. Si $\langle K,<\rangle$ es un conjunto bien ordenado y $f\colon K \to K$ es un morfismo, $x \leq f(x)$ para todo $x \in K$.

Indudablemente, $x \leq f(x)$ si $x$ es el primer elemento de $K$. Supongamos que $x$ no es el primer elemento de $K$ y que $y \leq f(y)$ para todo $y < x$ . Como $f$ es un morfismo, $f(y) < f(x)$. Por lo tanto, $f(x)$ está precedido por *todos* los predecesores de $x$, de modo que $x \leq f(x)$.

PROPOSICIÓN 2. Si $\langle K,<\rangle$ es un conjunto bien ordenado el único automorfismo $f{:}K \to K$ es la identidad, $f(x) = x$.

Sea $f{:}K \to K$ un automorfismo. Por la Proposición 1, $x \leq f(x)$ para todo

$x \in K$. La aplicación inversa $f^{-1}$ también es un automorfismo, de suerte que $z \leq f^{-1}(z)$ para todo $z \in K$. Reemplazando $z$ por $f(x)$ en esta desigualdad, tenemos que $f(x) \leq f^{-1}(f(x)) = x \leq f(x)$ para todo $x \in K$. Por lo tanto, $x = f(x)$ para todo $x \in K$.

Proposición 3. Si los conjuntos bien ordenados $\langle K_1,<_1\rangle$ y $\langle K_2,<_2\rangle$ son isomórficos, hay un solo isomorfismo $f$: $K_1 \to K_2$.

Sean $f$: $K_1 \to K_2$ y $h$: $K_1 \to K_2$, isomorfismos. Entonces $f^{-1}$: $K_2 \to K_1$ es un isomorfismo y $f^{-1}h$ es un automorfismo de $\langle K_1,<_1\rangle$. Por lo tanto $f^{-1}h$ es la identidad sobre $K_1$, y $f = h$.

Proposición 4. Si $\langle K,<\rangle$ es un conjunto bien ordenado y $a$ es cualquier elemento de $K$, $\langle K,<\rangle$ no es isomórfico a $\langle [a],<\rangle$, el segmento de $K$ determinado por $a$.

Para demostrar esta proposición, suponemos que es falsa y que hay un isomorfismo $f$ con dominio $K$ y alcance $[a]$. Por la Proposición 1, $a \leq f(a)$, ya que $f$: $K \to K$ es un morfismo. Pero esto contradice la suposición de que el alcance de $f$ es el conjunto de los predecesores de $a$.

Proposición 5. Si los conjuntos bien ordenados $\langle K_1,<_1\rangle$ y $\langle K_2,<_2\rangle$ no son isomórficos, entonces $\langle K_1,<_1\rangle$ es isomórfico a una segmento de $\langle K_2,<_2\rangle$ o $\langle K_2,<_2\rangle$ es isomórfico a un segmento de $\langle K_1,<_1\rangle$.

Definiré una aplicación $f_x$:$[x] \cup \{x\} \to K_2$ para cada $x \in K_1$ que cumpla una de estas dos condiciones: o bien $x$ es el primer elemento de $K_1$, o bien $f_z$ está definida para todo $z <_1 x$ y el alcance de $f_z$ no es igual a $K_2$. Doy la definición de $f_x$ mediante dos condiciones que corresponden, respectivamente, a cada una de las dos partes disjuntas mencionadas en la descripción del dominio de $f_x$, a saber, $\{x\}$ y $[x] = \{z \in K_1: z <_1 x\}$:

(1) Sea $u$ el primer elemento de $K_2$ tal que, para todo $z <_1 x$, $u \neq f_z(z)$. $f_x(x) = u$.

(2) Para todo $z <_1 x$, $f_x(z) = f_z(z)$.

Como es obvio, esta definición implica que si $x$ es el primer elemento de $\langle K_1,<_1\rangle$, $f_x(x)$ es el primer elemento de $\langle K_2,<_2\rangle$. Si $f_x$ está definida para cada $x \in K_1$, $x \mapsto f_x(x)$ es un isomorfismo de $\langle K_1,<_1\rangle$ en un segmento de $\langle K_2,<_2\rangle$

(su alcance no puede ser igual a $K_2$ ya que, por hipótesis, $\langle K_1,<_1\rangle$ y $\langle K_2,<_2\rangle$ no son isomórficos). En cambio, si $f_x$ no está definida para algún $x \in K_1$ y $t$ es el primer elemento de $K_1$ tal que $f_t$ no está definida, la aplicación $f$: $[t] \rightarrow K_2$; $x \mapsto f_x(x)$ es un isomorfismo de un segmento de $\langle K_1,<_1\rangle$ en $\langle K_2,<_2\rangle$.

Proposición 6: Si $\langle K,<\rangle$ es un conjunto bien ordenado y $H \subseteq K$, $\langle H,<\rangle$ es un conjunto bien ordenado isomórfico a $\langle K,<\rangle$ o a un segmento de $\langle K,<\rangle$.

Que $\langle H,<\rangle$ está bien ordenado se prueba así: si $H = \varnothing$, $\langle H,<\rangle$ está por cierto bien ordenado; si $H \neq \varnothing$ y $G$ es una parte no vacía de $H$, sea $G'$ la unión de los remanentes determinados en $K$ por cada elemento de $G$. Entonces, $G'$ es el remanente determinado en $K$ por algún $u \in G$, y $u$ precede a todos los otros elementos de $G$. Ahora bien, si $\langle K,<\rangle$ fuese isomórfico a un segmento $[u]$ de $\langle H,<\rangle$ habría un isomorfismo $f$: $K \rightarrow [u]$. En tal caso, como $u \in K$, $f(u) \in [u]$ y, por consiguiente, $f(u) < u$. Pero según la Proposición I esto no es posible si $f$ preserva el orden <. Por lo tanto, en virtud de la Proposición 5, $\langle H,<\rangle$ es isomórfico a $\langle K,<\rangle$ o a un segmento de $\langle K,<\rangle$.

## III

## LA CARDINALIDAD DE LA SEGUNDA CLASE DE ORDINALES

Aquí bosquejaré una demostración de que (i) el conjunto de los ordinales infinitos de la segunda clase es más numeroso que el conjunto $[\omega]$ de los ordinales finitos y que (ii) no existe una numerosidad intermedia entre la de este conjunto y la de aquél. Con la nomenclatura utilizada en el Capitulo 1.5, estas dos tesis pueden expresarse así: (i) $\aleph_0 \prec |[\omega_1]\setminus[\omega]|$ y (ii) no hay un conjunto de ordinales $K$, tal que $\aleph_0 \prec |K| \prec |[\omega_1]\setminus[\omega]|$. Ello implica que $|[\omega_1]\setminus[\omega]| = \aleph_1$. Los argumentos que aduciré para establecer este resultado provienen de Cantor 1897 (GA, pp. 331–33).

(i) Sea $n \mapsto \alpha_n$ una aplicación inyectiva de $[\omega]$ en la Clase (II). Si la lista $\alpha_0, \alpha_1, \alpha_2, \ldots$ contiene todos los ordinales de la Clase (II), éstos tienen que

figurar en ella en un orden distinto del fijado por la relación <.[2] Sea $k_0$ el índice más bajo asignado en esta lista a un sucesor de $\alpha_0$ (esto es: $\alpha_0 < \alpha_{k_0}$, y $\alpha_{k_0} < \alpha_k$ sólo si $k_0 \leq k$). Sea $k_{n+1}$ el índice más bajo asignado en la lista a un sucesor de $\alpha_{k_n}$ ($n \geq 0$). Entonces, la lista $\alpha_0, \alpha_{k_0}, \alpha_{k_1}, \ldots$ es una secuencia infinita, ordenada por la relación <, de números de la Clase (II) ninguno de los cuales sucede a todos los otros. Por lo tanto, existe un ordinal $\alpha$, de la Clase (II), que es el sucesor próximo de los ordinales contenidos en la lista $\alpha_0, \alpha_{k_0}, \alpha_{k_1}, \ldots$ En otras palabras, existe un $\alpha$ tal que $\alpha_{k_n} < \alpha$ para todo $n \geq 0$ y $\alpha$ es menor que cualquier otro ordinal —si lo hubiere— que sea mayor que todos los ordinales en la lista $\alpha_0, \alpha_{k_0}, \alpha_{k_1}, \ldots$ Si $\alpha$ figurara en la lista $\alpha_0, \alpha_1, \alpha_2, \ldots$ tendríamos que $\alpha = \alpha_k \leq \alpha_{k_n}$ para algún $k$ y $n \geq 0$. Por lo tanto, $\alpha$ no pertenece al alcance de la aplicación $n \mapsto \alpha_n$. En consecuencia, dicha aplicación no es biyectiva. Este razonamiento vale para cualquier inyección de $[\omega]$ en la Clase (II). Por lo tanto, la Clase (II) es más numerosa que $[\omega]$.

(ii) Sea $K$ un conjunto infinito de ordinales de la Clase (II). Sea $K' = K \cup [\omega]$. Entonces, por la Proposición 6 del Apéndice III, $\langle K', < \rangle$ es un conjunto bien ordenado isomórfico a $\langle [\omega_1], < \rangle$ o a un segmento $[\alpha]$ de $\langle [\omega_1], < \rangle$, donde $\alpha$ es un ordinal de la Clase (II). En el primer caso, $|K'| = |[\omega_1]|$ y $|K| = |[\omega_1] \setminus [\omega]|$. En el segundo caso, $|K| \preccurlyeq |K'| = |[\alpha]| = \aleph_0$. Por lo tanto, no hay un conjunto de ordinales de la Clase (II) que sea menos numeroso que la Clase (II) pero más numeroso que la Clase (I).

2 En el orden fijado por la relación < hay muchos ordinales de la Clase II que no tienen sucesor inmediato; pero en el orden de la lista $\alpha_1, \alpha_2, \ldots$, cada ordinal $\alpha_n$ tiene un sucesor inmediato $\alpha_{n+1}$.

# IV

## EL ARGUMENTO DE BURALI-FORTI

Burali-Forti (1897) se propone “demostrar que hay *números transfinitos* (o *tipos de orden*) $a$ y $b$, tales que $a$ no es igual a $b$ ni menor que $b$ ni mayor que $b$”. El trabajo adolece de cierta desidia terminológica, patente ya en esta primera oración, que equipara las expresiones “número transfinito” y “tipo de orden” (con explícita referencia a Cantor 1895). En el vocabulario del grupo de Peano, al que pertenecía Burali-Forti, la palabra ‘clase’ (Fr. o It. ‘classe’) se usaba como sinónimo de ‘conjunto’ (“ensemble quelconque d’objets”—Peano 1894, p. 4). Burali-Forti emplea ‘orden’ como sinónimo de lo que aquí he llamado ‘orden lineal’ y define un *tipo de orden* como una clase de clases ordenadas, equivalentes por isomorfismo. Por una inadvertencia (corregida en Burali-Forti 1897a), Burali-Forti entiende que una clase bien ordenada en el sentido de Cantor es una clase ordenada que cumple las condiciones BO1 y BO2 del Apéndice I. Por su parte, define una *clase perfectamente ordenada* como una clase ordenada $\langle M,<\rangle$ que reúne las condiciones BO1 y BO2 y llena además el siguiente requisito:

BF Si $x \in M$ y $x$ tiene un predecesor inmediato,[3] hay un $u_0 \in M$ tal que $u_0 < x$ y una clase finita $\{u_0, u_1, \ldots, u_n\} \subseteq M$, tales que $u_k$ es el predecesor inmediato de $u_{k+1}$ $(0 \leq k \leq n)$ y $u_n$ es el predecesor inmediato de $x$.

BF se deduce de la condición BO que define un conjunto bien ordenado en el genuino sentido cantoriano, pero no equivale a ella. Por lo tanto, si $M$ está bien ordenado en el sentido de Cantor, está perfectamente ordenado en el sentido de Burali-Forti, pero no a la inversa. Burali-Forti define un ‘número ordinal’ como el tipo de orden de un conjunto perfectamente ordenado. Evidentemente, habrá ordinales de Burali-Forti que no sean ordinales

[3] Si $<$ determina un orden lineal en $M$, digo que $y \in M$ es el predecesor inmediato de $x \in M$ si $y < x$ y, cualquiera que sea $z \in M$, $z < x \Rightarrow z \leq y$. La expresión ‘sucesor inmediato’ se define reemplazando ‘<’ por ‘>’ y ‘≤’ por ‘≥’ en la definición anterior.

cantorianos. En el resto de este apéndice usaré minúsculas griegas para designar ordinales de Burali-Forti. Designaré con $O$ a la clase $\{\xi$: $\xi$ es un ordinal de Burali-Forti$\}$. Sean $\langle A,R \rangle$ y $\langle B,S \rangle$ dos clases perfectamente ordenadas, con tipo de orden $\alpha$ y $\beta$, respectivamente. Burali-Forti dice que $\alpha < \beta$ ('$\alpha$ es menor que $\beta$') si $\langle A,R \rangle$ es isomórfica a una parte de $\langle B,S \rangle$ y $\langle B,S \rangle$ no es isomórfica a una parte de $\langle A,R \rangle$. La relación $<$ es irreflexiva (si $\alpha < \beta$, es falso que $\alpha = \beta$), antisimétrica (si $\alpha < \beta$, es falso que $\beta < \alpha$) y transitiva (si $\alpha < \beta$ y $\beta < \gamma$, $\alpha < \gamma$), pero $<$ no ordena a la clase $O$ debido a que, según el teorema anunciado al principio, hay ordinales $\alpha$ y $\beta$ tales que ni $\alpha < \beta$, ni $\alpha = \beta$, ni $\beta < \alpha$.

La prueba del teorema depende de dos proposiciones, $P_1$ y $P_2$, que demostraré luego: ($P_1$) Si $\zeta$ es un ordinal, existe un ordinal $\zeta'$ mayor que $\zeta$ tal que, si $\xi$ es un ordinal mayor que $\zeta$, $\zeta < \zeta' \leq \xi$ ($\zeta'$ es el sucesor inmediato de $\zeta$). ($P_2$) Si $\langle O,< \rangle$ es una clase perfectamente ordenada, con tipo de orden $\Omega \in O$, y $\xi$ es cualquier ordinal, $\xi \leq \Omega$. Burali-Forti argumenta así: Supongamos que $<$ ordena a $O$. Entonces $\langle O,< \rangle$ es una clase perfectamente ordenada,[4] y su tipo de orden $\Omega \in O$. En virtud de $P_1$ existe un ordinal $\Omega' > \Omega$. En virtud de $P_2$, $\Omega' \leq \Omega$. Esta contradicción implica que $<$ no ordena a $O$ y que, por lo tanto, hay ordinales $\alpha$ y $\beta$ tales que ni $\alpha < \beta$, ni $\alpha = \beta$, ni $\beta < \alpha$.

Para establecer la proposición $P_1$, consíderese cualquier clase perfectamente ordenada $\langle A,R \rangle$ con tipo de orden $\alpha$ y un objeto $z \notin A$. En la clase $A \cup \{z\}$ definimos una relación de orden $T$ como sigue: $xTy$ si $xRy$ o si $x \in A$ e $y = z$. $\langle A \cup \{z\},T \rangle$ es una clase perfectamente ordenada con tipo de orden $\alpha' > \alpha$.

4 $\langle O,< \rangle$ cumple las condiciones BO1 y BO2, puesto que hay un primer ordinal y cada ordinal tiene un sucesor inmediato (por $P_1$). Para ver que $\langle O,< \rangle$ cumple además la condición BF recuérdese que si $\alpha$ es un ordinal, $\alpha$ es el tipo de orden de una clase perfectamente ordenada $\langle M,P \rangle$. Como $\langle M,P \rangle$ cumple la condición BF, si $\alpha$ tiene un predecesor inmediato en $\langle O,< \rangle$, hay un $u \in M$ tal que el predecesor inmediato de $\alpha$ es el tipo de orden de la clase perfectamente ordenada $\langle M \setminus \{u\},P \rangle$. $u$ tiene un predecesor inmediato en $\langle M,P \rangle$ y en virtud de la condición BF, $M$ contiene una subclase finita $\{u_0, u_1, \ldots, u_n = u\}$, tal que $u_0$ no tiene predecesor inmediato y, para cada entero $k$ ($0 \leq k \leq n$), $u_k$ es el predecesor inmediato de $u_{k+1}$. $\langle M \setminus \{u_k, u_{k+1}, \ldots, u_n\},P \rangle$ es una clase perfectamente ordenada. Sea $\beta_k$ su tipo de orden. Vemos entonces que $O$ contiene una subclase finita $\{\beta_0, \beta_1, \ldots, \beta_n = \alpha\}$, tal que $\beta_0$ no tiene predecesor inmediato y para cada entero $k$ ($0 \leq k \leq n$), $\beta_k$ es el predecesor inmediato de $\beta_{k+1}$. Por lo tanto, $\langle O,< \rangle$ cumple la condición BF.

En vista del modo como se ha construido $\alpha'$, es claro que no puede haber un ordinal $\xi$ mayor que $\alpha$ y menor que $\alpha'$. Esta demostración de $P_1$ no sería aplicable a un ordinal $\alpha$ si $\alpha$ fuera el tipo de orden de la clase perfectamente ordenada de todo lo que hay (si hubiera tal clase), pues entonces no habría un objeto $z$ que pudiera agregarse a dicha clase para construir otra con tipo de orden $\alpha' > \alpha$. Pero en la demostración del teorema se considera sólo el tipo de orden de $\langle O,<\rangle$, y hay por cierto muchos objetos que no son ordinales de Burali-Forti. Si $q \notin O$, es posible extender de una manera obvia el orden $<$ a la clase $O \cup \{q\}$. Entonces $\langle O \cup \{q\},<\rangle$ tiene un tipo de orden $\kappa$ tal que $\kappa \in O$ y $\kappa > \Omega$.

La proposición $P_2$ se establece así: Sea $\alpha$ el tipo de orden de la clase perfectamente ordenada $\langle A,R\rangle$ y sea $A_x = \{z : z \in A \wedge \neg xRz\}$, para cualquier $x \in A$. Entonces, la clase $W = \{A_x: x \in A\}$ es una clase de clases perfectamente ordenadas, cada uno de cuyos elementos es una subclase de $A$. Digamos que $A_xSA_y$ si y sólo si $xRy$. $W$ está ordenada por la relación $S$ y las clases ordenadas $\langle W,S\rangle$ y $\langle A,R\rangle$ son isomórficas. Por lo tanto, $\langle W,S\rangle$ es una clase perfectamente ordenada con tipo de orden $\alpha$. Sea $[\alpha]$ la clase de ordinales $\{\alpha_x: x \in A$ y $\alpha_x$ es el tipo de orden de $\langle A_x,S\rangle\}$. $[\alpha]$ reune todos los ordinales menores que $\alpha$. Evidentemente, $\langle[\alpha],<\rangle$ es una clase ordenada isomórfica a $\langle W,S\rangle$. Por lo tanto, el tipo de orden de $\langle[\alpha],<\rangle$ es $\alpha$. Como $\langle[\alpha],<\rangle$ es isomórfica a una parte de $\langle O,<\rangle$, es claro que $\alpha \leq \Omega$.

Como ya habrá advertido el lector, la contradicción generada por Burali-Forti para probar su teorema no tiene nada de paradójica. Como en cualquier *reductio ad absurdum*, la contradicción sirve eficazmente para refutar la negación de la tesis que el autor se ha propuesto demostrar. El teorema puede probarse también sin mencionar el tipo de orden $\Omega$ de la clase de los ordinales de Burali-Forti. Basta tener presente que hay ordinales de Burali-Forti que no son ordinales de Cantor. Sea $\langle A,R\rangle$ un conjunto bien ordenado con tipo de orden $\alpha$ y sea $\langle B,S\rangle$ un conjunto perfectamente ordenado pero no bien ordenado, con tipo de orden $\beta$. Supongamos que $|B| \prec |A|$. En tal caso, $\langle A,R\rangle$ no puede ser isomórfico a una parte de $\langle B,S\rangle$, de modo que $\alpha \neq \beta$ y $\alpha$ no es menor que $\beta$. Sea $H \subseteq A$ tal que $|H| = |B|$ y sea $f{:}H \to B$ biyectiva. $f$ no puede ser un isomorfismo entre $\langle H,R\rangle$ y $\langle B,S\rangle$, por cuanto $\langle H,R\rangle$ es un conjunto bien ordenado (por la Proposición 6 del Apéndice II).[5]

[5] Como $\langle B,S\rangle$, por hipótesis, no es un conjunto bien ordenado, hay un subconjunto no vacío $K \subset B$ que no tiene un primer elemento (en el orden $S$). El conjunto $f^{-1}K$ de los

El argumento de Burali-Forti es aplicable también a los ordinales cantorianos, esto es, los tipos de orden de los conjuntos bien ordenados. Pero en este caso, tenemos también una prueba —basada en la Proposición 5 del Apéndice II— de que la clase de los ordinales está bien ordenada por $<$. Esta prueba depende esencialmente de la referencia al ordinal cantoriano de la clase de todos los ordinales cantorianos y constituye una refutación de la hipótesis de que dicho ordinal existe. En su correspondencia con Dedekind, Cantor indica que su inexistencia se debe a que los ordinales no se dejan reunir todos en un conjunto (Cantor, GA, pp. 443ss.).

## V

## LA NUEVA DEMOSTRACION DEL TEOREMA DEL BUEN ORDEN (ZERMELO 1908)

Sea $M$ un conjunto no vacío y $f$ un selector para $\mathcal{P}M\setminus\{\varnothing\}$. Llamo $A_{\setminus f}$ al conjunto $A\setminus\{f(\mathrm{A})\}$. (Zermelo lo llama $A'$, pero en la nomenclatura ahora habitual esta expresión designa a $M\setminus A$, el complemento de $A$ en $M$). Una $f$-cadena de $M$ es un conjunto $K \subseteq \mathcal{P}M$ tal que:

C1. $M \in K$.
C2. Si $A \in K$, $A_{\setminus f} \in K$.
C3. Si $H \subseteq K$, $\bigcap H \in K$.

Esta definición implica que la intersección de varias $f$-cadenas de $M$ también es una $f$-cadena en $M$. Con Zermelo, llamaré $\mathsf{M}$ a la intersección de

elementos de $H$ que $f$ envía a $K$ es una parte del conjunto bien ordenado $\langle H,R\rangle$ y por lo tanto tiene un primer elemento $u$ (en el orden $S$). Pero $f(u)$ no puede ser el primer elemento de $f^{-1}K$ y por lo tanto hay un $v \in K$ tal que $vSf(u)$. $f^{-1}(v) \in f^{-1}K$ y si $f$ fuera un isomorfismo, tendríamos que $f^{-1}(v)Ru$, contraviniendo el supuesto de que $u$ es el primer elemento de $f^{-1}K$.

todas las $f$-cadenas de $M$.[6] El lector debe comprobar que $\mathsf{M} \subseteq \mathcal{P}M$ cumple las condiciones C1–C3.

El primer paso de la demostración consiste en construir un conjunto $\Delta \subseteq \mathsf{M}$ con propiedades interesantes para la tarea entre manos. $A \in \Delta$ si y sólo si $A \in \mathsf{M}$ y, para todo $X \in \mathsf{M}$, $X \neq A$ implica que, o bien $X \subseteq A$, o bien $A \subseteq X$. Mostraremos que $\Delta$ es una $f$-cadena en $M$. Como $\mathsf{M}$ es la intersección de todas las $f$-cadenas de $M$, esto implica que $\mathsf{M} \subseteq \Delta \subseteq \mathsf{M}$, es decir, que $\Delta = \mathsf{M}$.

$\Delta$ cumple la condición C1, puesto que, evidentemente, $M \in \Delta$.

Si $A \in \Delta$, llamo $\mathcal{U}_A$ al conjunto $\{X \in \mathsf{M}: X \neq A \wedge X \subseteq A\}$ y $\mathcal{V}_A$ al conjunto $\{X \in \mathsf{M}: X \neq A \wedge A \subseteq X\}$. En virtud de la definición de $\Delta$, tenemos que $M = \mathcal{U}_A \cup \{A\} \cup \mathcal{V}_A$. Sea $\mathcal{W}_A = \{X \in \mathsf{M}: X \subseteq A_{\setminus f}\}$. Demostraré que $\mathcal{W}_A = \mathcal{U}_A$. Obsérvese que si $V \in \mathcal{V}_A$, $V \setminus A \neq \varnothing$ y que si $U \in \mathcal{U}_A$, $A \setminus U \neq \varnothing$. Ello implica que si $V \in \mathcal{V}_A$, $V_{\setminus f} \notin \mathcal{U}_A$. En efecto, o bien $f(V) \in A$, en cuyo caso $V \setminus A \subseteq V_{\setminus f}$ y $V_{\setminus f}$ contiene elementos que no pertenecen a $A$ ni, por ende, a ningún $U \in \mathcal{U}_A$, o bien $f(V) \notin A$, en cuyo caso $A \setminus U \subseteq V_{\setminus f}$ para cada $U \in \mathcal{U}_A$. Como $V_{\setminus f} \in \mathsf{M}$ (por C2), tenemos que $V_{\setminus f} \in \{A\} \cup \mathcal{V}_A$. Por la misma razón, si $\mathcal{F} \subseteq \mathcal{V}_A$, $\bigcap\mathcal{F} \in \{A\} \cup \mathcal{V}_A$. Por otra parte, $A_{\setminus f} \in \mathcal{W}_A$, y si $\mathcal{F} \subseteq \mathsf{M}$, $\bigcap\mathcal{F} \cap W \in \mathcal{W}_A$. Por consiguiente, $\mathcal{W}_A \cup \{A\} \cup \mathcal{V}_A$ es una $f$-cadena en $M$, de suerte que $\mathsf{M} \subseteq (\mathcal{W}_A \cup \{A\} \cup \mathcal{V}_A) \subseteq \mathsf{M}$. Por lo tanto, $(\mathcal{W}_A \cup \{A\} \cup \mathcal{V}_A) = \mathsf{M} = (\mathcal{U}_A \cup \{A\} \cup \mathcal{V}_A)$. Si $X \in M$, $X \in \mathcal{W}_A \Leftrightarrow X \notin \{A\} \cup \mathcal{V}_A \Leftrightarrow X \in \mathcal{U}_A$. Por lo tanto, $\mathcal{U}_A = \mathcal{W}_A$.

Este resultado nos permitirá probar que para todo $A \in \Delta$, $A_{\setminus f} \in \Delta$, de modo que $\Delta$ cumple la condición C2. Sean $\mathcal{U}_A{}^* = \{X \in \mathsf{M}: X \neq A_{\setminus f} \wedge X \subseteq A_{\setminus f}\}$ y $\mathcal{V}_A{}^* = \{X \in \mathsf{M}: X \neq A_{\setminus f} \wedge A_{\setminus f} \subseteq X\}$. Obviamente $\mathcal{V}_A{}^* = \{A\} \cup \mathcal{V}_A$ y $\mathcal{U}_A = \mathcal{U}_A \cup \{A_{\setminus f}\}$. Por lo tanto, $\mathsf{M} = \mathcal{U}_A{}^* \cup \{A_{\setminus f}\} \cup \mathcal{V}_A{}^*$, y $A_{\setminus f} \in \Delta$.

Para comprobar que $\Delta$ cumple también la condición C3, considérese un conjunto cualquiera $\Gamma \subseteq \Delta$. Si $X \in \mathsf{M} \setminus \{\bigcap\Gamma\}$, o bien existe un $Y \in \Gamma$ tal que $\bigcap\Gamma \subseteq Y \subseteq \mathrm{X}$, o bien $X \subseteq Y$ para todo $Y \in \Gamma$ y por ende $X \subseteq \bigcap\Gamma$. Esto indica que $\bigcap\Gamma \in \Delta$, como exije C3.

6 Esta nomenclatura no está exenta de peligro. En la p. 109, línea 10 del texto original, el tipógrafo de *Mathematische Annalen* escribió '$\mathsf{M}$' en lugar de '$M$'. El mismo error se repite, triplicado, en la traducción inglesa de van Heijenoort. Espero que el invento de la computadora de mesa, que me permite componer yo mismo este libro para la imprenta a medida que lo voy redactando, protegerá al lector contra este riesgo.

$\Delta$ cumple, pues, todos los requisitos para ser una $f$-cadena de $M$. Así queda demostrado que $\Delta = \mathsf{M}$. Escribo '$Y \supset X$' para decir que $Y$ incluye a $X$ como una parte propia ($X \subseteq Y \wedge X \neq Y$). En virtud de la misma definición de $\Delta$, es claro que, si $A$ y $B$ son dos elementos distintos de $\Delta = \mathsf{M}$, o bien $A \supset B$, o bien $A \supset B$. $\mathsf{M}$ está, pues, ordenado linealmente por la relación $\supset$. $\langle \mathsf{M}, \supset \rangle$ es un conjunto bien ordenado, puesto que cumple las condiciones BO1, BO2 y BO3 del Apéndice I. En efecto, $\langle \mathsf{M}, \supset \rangle$ tiene un primer elemento, a saber, $M$; cada $H \in \mathsf{M}$ tiene un sucesor inmediato, a saber $H_{\backslash f}$; por último, si $\mathscr{F} \subset \mathsf{M}$ tiene sucesores en $\langle \mathsf{M}, \supset \rangle$ (vale decir, si $\varnothing \notin \mathscr{F}$), $\bigcap \mathscr{F} \in \mathsf{M}$ es el sucesor próximo de $\mathscr{F}$.

Para demostrar que $f$ aplica $\mathsf{M}$ biyectivamente sobre $M$, Zermelo usa un argumento muy ingenioso. Muestra primero que si $\varnothing \neq P \subseteq M$, hay un y sólo un conjunto $P_0 \in \mathsf{M}$ tal que $f(P_0) \in P \subseteq P_0$. Reemplazando $P$ por $\{k\}$ en este resultado preliminar, se sigue que, para cada $k \in M$ hay un único $K_0 \in \mathsf{M}$ tal que $f(K_0) \in \{k\} \subseteq K_0$, esto es, tal que $f(K_0) = k$. El resultado preliminar se establece así: Sea $\varnothing \neq P \subseteq M$ y $P_0 = \bigcap \{X \in \mathsf{M} : P \subseteq X\}$. $P_0 \in \mathsf{M}$ (por C3) y $f(P_0) \in P$, ya que, si $f(P_0) \notin P$, $P \subseteq P_0 \setminus \{f(P_0)\}$ y $P_0$ no es la intersección de todos los elementos de $\mathsf{M}$ que incluyen a $P$. Por otra parte, si $P \subseteq P_1 \in \mathsf{M}$ y $P_1 \neq P_0$, $P_1 \supset P_0$, de modo que $f(P_1) \notin P_0$. Como $P \subseteq P_0$, $f(P_1) \notin P$. Así, $P_0$ resulta ser el único $X \in \mathsf{M}$ tal que $f(X) \in P \subseteq X$.

Como el selector $f$ restringido a $\mathsf{M}$ es biyectivo, induce en $M$ el buen orden de $\langle \mathsf{M}, \supset \rangle$, según expliqué en el Capitulo 1.7. Zermelo prueba además que el buen orden determinado de este modo en un dado conjunto $M$ por un dado selector $f: \mathscr{P}M \to M$ es *único*. Supongamos, en efecto, que $\langle M, < \rangle$ es un conjunto bien ordenado de tal modo que, para cada $A \subseteq M$, $f(A)$ es el primer elemento de $\langle A, < \rangle$. Entonces, cada $k \in M$ determina una parte de $M$ formada por $k$ y sus sucesores. Sea $H_k = \{x \in M: k \leq x\}$. Es claro que $k = f(H_k)$. Sea $\mathsf{H} = \{H_k: k \in M\}$. Para cada $P \subseteq M$ hay un y sólo un $H_P \in \mathsf{H}$ tal que $f(P) = f(H_P)$. $H_P$ contiene el primer elemento de $P$ y *todos* sus sucesores y, por ende, cumple la condición $f(H_P) \in P \subseteq H_P$. No es difícil comprobar que $\mathsf{H}$ cumple las condiciones C1, C2 y C3. En otras palabras, $\mathsf{H}$ es una $f$-cadena de $M$. Sea $\mathsf{M}$, como antes, la intersección de todas las $f$-cadenas de $M$. Sabemos que, si $Z \in \mathsf{H}$, hay un y sólo un $U \in \mathsf{M}$ tal que $f(U) \in Z \subseteq U$. Como $\mathsf{M} \subseteq \mathsf{H}$ y $Z$ mismo es el único $X \in \mathsf{H}$ tal que $f(X) \in Z \subseteq \mathrm{X}$, tenemos que $\mathrm{U} = \mathrm{Z}$. Por lo tanto, $\mathsf{H} = \mathsf{M}$ y el orden $<$ es precisamente el buen orden inducido en $M$ por el selector $f$ restringido a $\mathsf{M}$.

# VI

## LOS AXIOMAS DE ZERMELO

Doy enseguida una traducción de los siete axiomas para la teoría de conjuntos formulados por Ernst Zermelo (1908b). En el artículo original el enunciado de los axiomas está enmarcado en explicaciones que en lo esencial he reproducido en la Sección 1.8.1. Allí traduje la definición del término *definit* utilizado en el Axioma III. Es oportuno advertir que Zermelo escribe ⋹ en vez de $\subseteq$, 0 en vez de $\varnothing$, $\mathfrak{U}T$ en vez de $\mathscr{P}T$, y $\mathfrak{S}T$ en vez de $\bigcup T$.

AXIOMA I Si cada elemento de un conjunto $M$ es a la vez un elemento de un conjunto $N$ y vice versa, o sea, si $M \subseteq N$ y $N \subseteq M$, entonces $M = N$. En suma, cada conjunto está determinado por sus elementos. (*Axioma de Determinación*).

AXIOMA II Hay un conjunto (impropio), el conjunto vacío $\varnothing$, que no contiene ningún elemento. Si $a$ es una cosa cualquiera del dominio, existe un conjunto $\{a\}$ que contiene a $a$ y sólo a $a$ como elemento. Si $a$ y $b$ son dos cosas cualesquiera del dominio, existe un conjunto $\{a, b\}$ que contiene como elementos a $a$ y $b$, pero no a otra cosa diferente $x$. (*Axioma de los conjuntos elementales*).

AXIOMA III Si la proposición universal $\mathfrak{E}(x)$ está bien definida (*definit*) para todos los elementos de un conjunto $M$, $M$ posee siempre un subconjunto $M_{\mathfrak{E}}$ que contiene como elementos a todos los elementos $x$ de $M$ para los cuales $\mathfrak{E}(x)$ es verdadera, y sólo a ellos. (*Axioma de Separación*).

AXIOMA IV A cada conjunto $T$ le corresponde un segundo conjunto $\mathscr{P}T$ (el *conjunto potencia* de $T$), que contiene como elementos a todos los subconjuntos de $T$, y sólo a ellos. (*Axioma del Conjunto Potencia*).

AXIOMA V A cada conjunto $T$ le corresponde un conjunto $\bigcup T$ (el *conjunto unión* de $T$), que contiene como elementos a todos los elementos de los elementos de $T$, y sólo a ellos. (*Axioma de Unión*).

AXIOMA VI Si $T$ es un conjunto todos cuyos elementos son conjuntos no vacíos que no comparten elementos,[7] la unión $\bigcup T$ incluye[8] al menos un subconjunto $S_1$ que tiene con cada elemento de $T$ un y sólo un elemento en común. (*Axioma de Selección*).

AXIOMA VII El dominio contiene al menos un conjunto $Z$ que contiene como elemento al conjunto vacío y está constituido de tal modo que a cada elemento suyo $a$ le corresponde otro elemento de la forma $\{a\}$, o que junto con cada elemento $a$ contiene también como elemento el conjunto correspondiente $\{a\}$. (*Axioma del Infinito*).

## VII

## INDEPENDENCIA DEL AXIOMA DE SELECCIÓN (FRAENKEL 1922a)

Para demostrar que una proposición φ es independiente de un conjunto de proposiciones Γ se concibe una situación objetiva en la cual se cumplen las proposiciones de Γ, pero φ es falsa. Una situación objetiva en la que se cumplen todas las proposiciones de Γ se llama un *modelo* de Γ. Para demostrar que el Axioma VI de Zermelo es independiente de los Axiomas I–V y VII, Fraenkel (1922a) propone un modelo cuyo dominio $\mathfrak{D}$ consta exclusivamente de estos objetos: (i) el conjunto vacío $\varnothing$; (ii) una colección infinita

[7] La frase “no comparten elementos” corresponde a la expresión “untereinander elementenfremd sind” utilizada por Zermelo. Debe entenderse que los elementos de $T$ son mutuamente disjuntos por pares, esto es, que si $A$ y $B$ son elementos de $T$, o bien $A = B$, o bien $A \cap B = \varnothing$.

[8] Zermelo usa aquí, para expresar la relación entre un conjunto y sus subconjuntos, el mismo verbo alemán ‘enthalten’ con que expresa la relación entre un conjunto y sus elementos en los Axiomas II, III, IV, V y VII. Para distinguir las dos relaciones, traduzco ‘enthalten’ por ‘incluir’ en el presente pasaje y por ‘contener’ en los otros.

pero enumerable de objetos distintos, $a_1$, $\overline{a}_1$, $a_2$, $\overline{a}_2$, . . . , ninguno de los cuales es un conjunto; (iii) el conjunto $Z_0 = \{\varnothing, \{\varnothing\}, \{\{\varnothing\}\}, \ldots\}$ postulado por el Axioma VII; (iv) el conjunto $A = \{\{a_1, \overline{a}_1\}, \{a_2, \overline{a}_2\}, \ldots\}$; y (v) todos los conjuntos que existen en virtud de los Axiomas I–V si existen los objetos descritos en las cuatro cláusulas anteriores. Todos los objetos de $\mathscr{D}$, excepto los mencionados en la cláusula (ii), son conjuntos y pueden construirse a partir de los "objetos primitivos" mencionados en las cláusulas (i), (ii) y (iii) mediante un número finito de aplicaciones de los Axiomas I–V.

Para facilitar la exposición, Fraenkel introduce algunos términos. El conjunto $\{a_k, \overline{a}_k\}$ se llama la *celda* $A_k$, con *elementos conjugados* $a_k$ y $\overline{a}_k$ ($k \geq 1$). Un *conjunto principal* es un conjunto que difiere de $A$ a lo sumo por la ausencia de un número finito de celdas. Si $M$ es un conjunto cualquiera, el conjunto $\overline{M}^k$ conjugado con $M$ respecto de la celda $A_k$ es el conjunto que difiere de $M$ solamente en cuanto cada elemento de $A_k$ que entra en la construcción de $M$ ha sido sustituido por su elemento conjugado. Si $M = \overline{M}^k$, decimos que $M$ es *simétrico* respecto de $A_k$. Si el conjunto $M$ es simétrico respecto de todas las celdas de un conjunto principal diré que $M$ es *idóneo* (Fraenkel no se vale de este término). Es fácil ver que todos los objetos primitivos postulados en las cláusulas (i)–(iii) son idóneos: en efecto, cualquiera que sea el entero positivo $k$, un intercambio entre $a_k$ y $\overline{a}_k$ en nada afecta a los conjuntos $\varnothing$, $Z_0$ y $A$; por su parte, $a_k$ y $\overline{a}_k$ son obviamente simétricos respecto de todas las celdas del conjunto principal $A \setminus \{a_k, \overline{a}_k\}$. Diré asimismo que una función φ (en el sentido de Fraenkel 1922a explicado en la Sección 1.8.2) es *idónea* si existe un conjunto principal $B_\varphi$ tal que, para todo $k \geq 1$, si $A_k \in B_\varphi$, entonces, para todo $x$, $\varphi(\overline{x}_k)$ es el conjunto conjugado con $\varphi(x)$ respecto de $A_k$.

La falsedad del Axioma de Selección en el modelo propuesto es una consecuencia inmediata de la siguiente "proposición fundamental", verdadera en dicho modelo:

PF. Todo conjunto es idóneo.

En efecto, si el Axioma de Selección es verdadero, existe un conjunto $S$ que "selecciona" un y sólo un elemento de cada celda en el conjunto $A$ y evidentemente $S$ no puede ser simétrico respecto de todas las celdas de un conjunto principal.

La verdad de PF en el modelo de Fraenkel resulta de cinco lemas que me limito a enunciar:

I. Si *M* es un conjunto, también existe el conjunto conjugado $\overline{M}^k$ (para cada entero positivo *k*).

II. Si *M* es un conjunto construido a partir de conjuntos idóneos por aplicación de los Axiomas II, IV y V (exclusivamente), *M* es idóneo.

III. Sea φ una función que se define sustituyendo por una variable alguno de los conjuntos idóneos presupuestos en la construcción del conjunto *M* mencionado en el lema II. Entonces φ es una función idónea.

IV. Si φ y ψ son funciones como la descrita en el lema III y *M* es un conjunto idóneo, entonces también son idóneos los subconjuntos de *M* determinados por las condiciones $\varphi \in \psi$ y $\varphi \notin \psi$ conforme al Axioma III (en la versión de Fraenkel reproducida en la Sección 1.8.2).

V. Si en el enunciado de las condiciones del lema IV el conjunto *M* o uno de los argumentos de las funciones φ o ψ es reemplazado por una variable, la condición $\varphi \in \psi$ (o la condición $\varphi \notin \psi$) determina una función idónea χ.

No es difícil probar que los cinco lemas son verdaderos en el modelo considerado (de hecho, Fraenkel sólo estima necesario dar pruebas del II y el IV). PF se deriva luego por inducción sobre la "clase" de cada objeto o función de $\mathscr{D}$. Este concepto de "clase" se define recursivamente así:

[0] Un objeto de $\mathscr{D}$ es un objeto de clase 0 si es uno de los objetos primitivos mencionados en las cláusulas (i)–(iii) o se construye por aplicación de los Axiomas II, IV y V a objetos primitivos. Una función definida sustituyendo por una variable alguno de los objetos primitivos presupuestos en la construcción de un objeto de clase 0 es una función de clase 0.

[$n \Rightarrow n$+1] Si el conjunto *M* y las funciones φ y ψ son de clase $\leq n$ y $M'$ es el subconjunto de *M* determinado conforme al Axioma III por la condición $\varphi \in \psi$ o por la condición $\varphi \notin \psi$, $M'$ es un conjunto de clase

> $\leq n+1$. Si $\chi$ es la función determinada en tal caso por la condición $\varphi \in \psi$ o por la condición $\varphi \notin \psi$ cuando el conjunto $M$ o uno de los argumentos de las funciones $\varphi$ o $\psi$ es reemplazado por una variable, $\chi$ es una función de clase $\leq n+1$.

El argumento inductivo procede así: Todo objeto de clase $\leq 1$ es idóneo en virtud de los lemas I y IV y toda función $\varphi$ de clase $\leq 1$ es idónea en virtud de los lemas III y V. Supongamos que se ha establecido que todos los objetos y funciones de clase $\leq n$ son idóneos. Entonces el lema IV implica que todo objeto de clase $\leq n+1$ es idóneo y el lema V implica que toda función de clase $\leq n+1$ es idónea.

A propósito de esta prueba de independencia es oportuno hacer algunas observaciones. Ella demuestra que el Axioma de Selección es independiente de los demás axiomas del sistema de Zermelo (1908a), con el Axioma III ajustado a la formulación de Fraenkel 1922a (Sección 1.8.2). Pero no se ha demostrado que el Axioma de Selección sea independiente del restante sistema de Zermelo enriquecido con el Axioma de Reemplazo propuesto en Fraenkel 1922 (Sección 1.8.3). De hecho, Fraenkel pensó en un comienzo que el Axioma de Selección se podía *inferir* de dicho sistema enriquecido. Por otra parte, si 'función' se entiende en el sentido de Fraenkel 1922a, el Axioma de Reemplazo no fortalece el sistema original de los Axiomas I–V y VII de Zermelo sino que *se infiere de ellos* (von Neumann 1928). El modelo de los Axiomas I–V y VII utilizado en la prueba de independencia comprende, como hemos visto, infinitos objetos que no son conjuntos. Como indiqué en la p. 75, nota 5, tales objetos se conocen en la literatura matemática por su nombre alemán de *Urelemente*. Zermelo estimaba que la teoría de conjuntos debía postular su existencia para que la matemática se pudiera aplicar al mundo real, en el cual hay por cierto muchas cosas que no son conjuntos. Pero la matemática pura puede vivir muy bien sin *Urelemente*. Por lo demás, tampoco los necesita la matemática aplicada, si lo que hace es *representar* realidades naturales mediante objetos ideales que perfectamente pueden ser conjuntos. En 1963 Paul Cohen demostró, sin presuponer *Urelemente*, que el Axioma de Selección es independiente de los axiomas restantes de Zermelo suplementados con el Axioma de Reemplazo (en la versión fuerte de Skolem).

# VIII

## LA DEFINICIÓN POR INDUCCIÓN TRANSFINITA (VON NEUMANN 1928a)

La inducción matemática (finita) es el arma principal de la aritmética. Una *demostración inductiva* establece que *todos* los números naturales tienen una propiedad *P*, probando (*a*) que el número 0 tiene la propiedad *P* y (*b*) que, si un número cualquiera *n* tiene la propiedad *P*, también la tiene su sucesor inmediato $n+1$. En el penúltimo párrafo del Apéndice VII hay un ejemplo de este método de demostración. Una *definición por inducción* (o *definición recursiva*) determina la extensión *C* de un concepto aplicable a objetos de cierta categoría *K*, por "estratos", como sigue:[9]

(i) Se estipula que cierta parte de *K* constituye el estrato básico $C_0$.

(ii) Se fija una relación funcional[10] $\phi$ entre las partes de *K* y se estipula que si *n* es cualquier número natural y el estrato $C_n$ tiene la relación $\phi$ con una parte $\phi C_n$ de *K*, el estrato $C_{n+1} = C_n \cup \phi C_n$.

(iii) *C* se define como la unión de todos los estratos $C_n$.

Dedekind fue el primero en comprender que la definición por inducción sólo es admisible si la sucesión de estratos $C_n$ ($n \in \mathbb{N}$) queda unívocamente determinada con sólo fijar el estrato básico $C_0$ y la relación $\phi$. La unicidad de los estratos y, por ende, de su unión *C* está garantizada por el siguiente teorema, que Dedekind (1888, #126) demuestra inductivamente: Dado un triple $\langle W, w_0, \phi\rangle$, donde *W* es un dominio de objetos, $w_0$ es un objeto distinguido de *W* y $\phi$ es una aplicación de *W* en *W*, existe una sola aplicación $f$: $\mathbb{N} \rightarrow W$ tal que $f(0) = w_0$ y para cada $n \in \mathbb{N}$, $f(n+1) = \phi(f(n))$. Sea $W = \mathcal{P}K$

[9] Acabamos de ver una definición recursiva (de 'clase') en el Apéndice VII. Hay otras dos en las pp. 78 y 81. El Apéndice IX contiene definiciones recursivas de 'término', 'predicado', 'fórmula', 'verdad/falsedad' de una fórmula en una interpretación, etc. Conviene examinar dichos ejemplos a la luz de la explicación que doy aquí.

[10] Una relación diádica *R* se dice *funcional* si la conjunción $Rxy \wedge Rxz$ implica que $y = z$; en otras palabras, si la relación *R* admite un solo *terminus ad quem* para cada *terminus a quo*.

y $w_0 = C_0$. Entonces, debido a la existencia y unicidad de la aplicación $f$ descrita en el teorema, hay una y sólo una aplicación de $\mathbb{N}$ en $W$, $k \mapsto C_k$, tal que para cada $n \in \mathbb{N}$, $C_{n+1} = f(n) \cup C_n$. Esto asegura la unicidad de la unión $C = \bigcup_{k\in\mathbb{N}} C_k$.[11]

La demostración inductiva se puede extender al transfinito. Se demuestra por inducción *transfinita* que cada ordinal tiene una propiedad $P$, probando ($\alpha$) que el 0 tiene la propiedad $P$ y ($\beta$) que, si todos los ordinales menores que un ordinal cualquiera $\xi$ tienen la propiedad $P$, también la tiene $\xi$.

También hay un método de *definición por inducción transfinita*, que construye la extensión de un concepto $C$ por estratos $C_0$, $C_1$, $C_2$, . . . , $C_\omega$, $C_{\omega+1}$, . . . correspondientes a los ordinales. Parafraseando a Hausdorff (1914, p. 113), este género de definición se puede caracterizar así:

$C_\kappa$ está definido para todo ordinal $\kappa$ si

(i) $C_0$ está definido y
(ii) para cualquier ordinal $\alpha$, $C_\alpha$ queda definido mediante la definición de $C_\xi$ para todo $\xi < \alpha$.

Aunque este método de definición se utiliza implícitamente desde los albores de la teoría de conjuntos, von Neumann parece haber sido el primero que juzgó necesario justificarlo. En la exposición siguiente seguiré a von Neumann 1928c. Imitando al autor, me expreso en términos de una teoría de conjuntos "ingenua", pero fácilmente adaptable al sistema ZF. La caracterización precisa de la definición por inducción transfinita se facilita mucho gracias al concepto de ordinal de von Neumann 1922. Recordemos que un ordinal así concebido es un conjunto $\alpha$ bien ordenado por la relación $\in$, de modo que cada elemento de $\alpha$ contiene a todos sus predecesores; el sucesor de $\alpha$ es el conjunto $\alpha \cup \{\alpha\}$; el primer ordinal es $0 = \varnothing$ (Sección 1.8.4). Llamaré $\mathfrak{B}$ —con Zermelo— al dominio de la teoría de conjuntos y $\mathbf{\Omega}$ —con Cantor— al subdominio constituido por los ordinales (von Neumann lo llama *OZ*). Si $f$ es una aplicación de $\mathbf{\Omega}$, o de un segmento de $\mathbf{\Omega}$, en $\mathfrak{B}$ y $\alpha$ es un ordinal, llamaré con von Neumann $F(f,\alpha)$ al grafo de la restricción de $f$ a $\alpha$; en otras palabras, $F(f,\alpha) = \{\langle\xi,f(\xi)\rangle: \xi \in \alpha\}$. Von Neumann demues-

11 En el Apéndice XI doy una demostración y ejemplos de aplicaciones del Teorema de la Definición por Inducción de Dedekind.

tra el siguiente Teorema de la Definición por Inducción Transfinita (TDIT), visiblemente similar al teorema de Dedekind (1888, #126) enunciado arriba:

> Si $\varphi$ es una aplicación definida en $\mathfrak{V} \times \mathfrak{V}$ hay una y sólo una aplicación $f: \mathbf{\Omega} \to \mathfrak{V}$ tal que para cada ordinal $\xi$, $f(\xi) = \varphi(F(f,\xi),\xi)$.

No es difícil ver que el TDIT legitima la definición por inducción transfinita. Recordemos que cada estrato $C_\xi$ ($\xi > 0$) de la extensión $C$ de un concepto definido por este método queda definido por la colección $\{C_\zeta: \zeta \in \xi\}$ de los estratos anteriores. La regla que coordina cada estrato con los que le preceden puede representarse, pues, mediante una aplicación $\varphi: \mathfrak{V} \times \mathfrak{V} \to \mathfrak{V}$, tal que $\varphi(0,0) = C_0$ y, para cada ordinal $\xi > 0$, $C_\xi = \varphi(\{\langle\zeta,C_\zeta\rangle: \zeta \in \xi\},\xi)$.[12] En virtud del TDIT, la aplicación $\xi \mapsto C_\xi$ existe y es única. El *definiendum* $C$ puede entonces identificarse con la unión de los estratos $C_\xi$: $C = \bigcup_{\xi\in\mathbf{\Omega}} C_\xi$.

Ahora bosquejaré la demostración de von Neumann. Diremos que un ordinal $\xi$ es *normal* si, dada una aplicación $\varphi: \mathfrak{V} \times \mathfrak{V} \to \mathfrak{V}$, existe una aplicación $f_\xi: \xi \to \mathfrak{V}$ tal que para cada $\zeta \in \xi$, $f_\xi(\zeta) = \varphi(F(f_\xi,\zeta),\zeta)$. Llamaré, con von Neumann, "elemento funcional hasta $\xi$" a una aplicación $f_\xi$ como la descrita. Von Neumann demuestra (I) que dados $\varphi$ y $\xi$ no puede existir más que un elemento funcional hasta $\xi$. Por lo tanto, si $\xi$ es normal, la expresión $\varphi(F(f_\xi,\xi),\xi)$ designa exactamente un objeto del dominio $\mathfrak{V}$.[13] Así pues, la correspondencia $\xi \mapsto \varphi(F(f_\xi,\xi),\xi)$ asigna un y sólo un objeto $f(\xi)$ a cada ordinal normal $\xi$. Von Neumann demuestra (II) que todo ordinal es normal. Por lo tanto, dada la aplicación $\varphi$, existe una única aplicación $f: \mathbf{\Omega} \to \mathfrak{V}$ tal que para cada ordinal $\xi$, $f(\xi) = \varphi(F(f_\xi,\xi),\xi)$. Q.E.D.

La proposición (I) se prueba así: Sea $\xi$ un ordinal normal y $f_\xi$ y $g_\xi$ dos elementos funcionales hasta $\xi$. Es claro que si $f_\xi \neq g_\xi$ hay uno o más ordinales $\zeta$ mayores que 0 y menores que $\xi$ tales que $f_\xi(\zeta) \neq g_\xi(\zeta)$. Sea $\zeta_0$ el menor de estos ordinales. Como $f_\xi(\eta) = g_\xi(\eta)$ para todo ordinal $\eta \in \zeta_0$, tenemos que, contra lo que se había supuesto, $f_\xi(\zeta_0) = \varphi(F(f_\xi,\zeta_0),\zeta_0)$

12 Por cierto, podría representársela también, más sencillamente, mediante una aplicación $\psi: \mathfrak{V} \to \mathfrak{V}$ tal que $\psi(0) = C_0$ y, para cada ordinal $\xi > 0$, $C_\xi = \psi(C_\zeta: \zeta \in \xi)$. La representación arriba descrita fue elegida por von Neumann con vistas a la demostración del TDIT.

13 Recuérdese que $F(f_\xi,\xi) = \{\langle\zeta,f_\xi(\zeta)\rangle: \zeta \in \xi\}$, de suerte que esta expresión está bien definida si $f_\xi$ es una aplicación definida en $\xi$.

$= \varphi(\{\langle\eta,f_\xi(\eta)\rangle: \eta \in \zeta_0\},\zeta_0) = \varphi(\{\langle\eta,g_\xi(\eta)\rangle: \eta \in \zeta_0\},\zeta_0) = \varphi(F(g_\xi,\zeta_0),\zeta_0) = g_\xi(\zeta_0)$. Por lo tanto, $f_\xi = g_\xi$.

La proposición (II) se infiere casi inmediatamente de las proposiciones siguientes:

(III) Si $\xi$ es un ordinal normal y $\zeta \in \xi$, $\zeta$ es normal y $\varphi(F(f_\zeta,\zeta),\zeta) = \varphi(F(f_\xi,\zeta),\zeta)$;

(IV) Si todo ordinal $\zeta \in \xi$ es normal, $\xi$ es normal.

Para probar (III) suponemos que $\xi$ es normal y designamos con $f_\xi|\zeta$ la restricción de $f_\xi$ a $\zeta \in \xi$. Si $\eta$ es cualquier ordinal menor que $\zeta$, $f_\xi|\zeta(\eta) = f_\xi(\eta)= \varphi(F(f_\xi,\eta),\eta) = \varphi(F(f_\xi|\zeta,\eta),\eta)$. Por lo tanto, $\zeta$ es normal y $f_\zeta = f_\xi|\zeta$. Es claro, pues, que $\varphi(F(f_\zeta,\zeta),\zeta) = \varphi(F(f_\xi|\zeta,\zeta),\zeta) = \varphi(F(f_\xi,\zeta),\zeta)$.

(IV) se prueba así: Sea $\xi$ un ordinal tal que si $\zeta \in \xi$, $\zeta$ es normal. En virtud de los resultados ya obtenidos, existe entonces para cada $\zeta \in \xi$ un objeto único $f_\xi(\zeta) = \varphi(F(f_\zeta,\zeta),\zeta)$. $\zeta \mapsto f_\xi(\zeta)$ es, pues, una aplicación definida en $\xi$.[14] En particular, por P3, si $\eta \in \zeta \in \xi$, $f(\eta) = \varphi(F(f_\eta,\eta),\eta) = \varphi(F(f_\zeta,\eta),\eta)$, de suerte que para todo $\zeta \in \xi$, $F(f,\zeta) = F(f_\zeta,\zeta)$, lo cual implica a su vez que

$$f_\xi(\zeta) = \varphi(F(f_\zeta,\zeta),\zeta) = \varphi(F(f_\xi,\zeta),\zeta)$$

Como sabemos, esto significa que $\xi$ es normal.

Para probar (II) suponemos que es falsa y que existe un ordinal $\alpha$ que no es normal. Entonces, por (IV), no son normales todos los ordinales $\xi$ tales que $\xi \in \alpha$. Von Neumann muestra que los ordinales no normales que son elementos de $\alpha$ constituyen un conjunto $\mathscr{A} \neq \varnothing$, el cual tiene, por lo tanto, un primer elemento $\xi_0$. Si $\zeta \in \xi_0$, $\zeta \in \alpha \setminus \mathscr{A}$; luego, $\zeta$ es normal. Pero entonces (IV) implica que $\xi_0$ es normal. ¡Contradicción!

[14] Von Neumann dedica una larga nota a demostrar, en términos adaptables a la teoría axiomática ZF, que la aplicación $\xi \to$ B definida por $\zeta \mapsto f_\xi(\zeta) = \varphi(F(f_\zeta,\zeta),\zeta)$ efectivamente existe si todo $\zeta < \xi$ es normal (1928a, pp. 389s., n. 28). Pero a la luz de lo que llevamos dicho su existencia debiera resultar obvia desde un punto de vista "ingenuo".

# IX

# EL CÁLCULO PREDICATIVO

*Cálculo predicativo* es el nombre genérico de los lenguajes artificiales o "escrituras conceptuales" (cf. Capítulo 2.2) utilizados en la lógica y la matemática actuales. Distinguimos entre cálculos de primer orden, en los que es posible representar formalmente generalizaciones relativas a objetos individuales (vgr. 'hay una montaña en América que es más alta que todas las montañas de Europa'), y cálculos de orden superior, en los que además pueden representarse generalizaciones relativas a propiedades y relaciones (vgr. 'hay una relación *X* que cualquier entero positivo *z* mayor que 3 tiene con algún número primo menor que *z*'). En los cálculos de segundo orden se representan generalizaciones relativas a individuos y a propiedades y relaciones de individuos; en los de tercer orden, se representan asímismo generalizaciones relativas a propiedades y relaciones de tales propiedades o relaciones, etc.

A continuación presento la gramática —sintaxis y semántica— común a toda una familia de cálculos predicativos de primer orden. Para facilitar la comprensión, presento primero reglas para cálculos sin signo de identidad, que llamo versiones del CP1, y luego enuncio reglas adicionales para cálculos con signo de identidad, que llamo versiones del CP1=. Enseguida, bajo el título "Decisión y deducción", considero dos vías para determinar cuáles son las fórmulas del CP1= que según nuestras estipulaciones semánticas representan verdades lógicas. Por último, me refiero brevemente a una familia de cálculos predicativos de segundo orden con identidad, que llamo CP2=.

En el curso de la exposición, se definen diversas categorías sintácticas (*términos*, *fórmulas*, etc.). Generalmente, la definición es *recursiva*: se fija primero una *base* formada por uno o más objetos bien definidos pertenecientes a la categoría en cuestión; luego se describe un procedimiento bien determinado para construir nuevos objetos de esa categoría a partir de objetos ya reconocidos como tales. Se *sobreentiende* que todo objeto de la categoría así definida satisface uno de estos dos criterios de definición y que aquellos objetos que no pertenecen a la base se generan desde ella mediante un número *finito* de aplicaciones del procedimiento descrito. Consíderese, por

ejemplo, la definición de *variable* bajo el acápite *Vocabulario* en la sección A: el objeto $x$ es una variable (éste es el único elemento de la base); si un objeto dado $\xi$ es una variable, se forma una nueva variable agregando el palote $\shortmid$ a la derecha de $\xi$; por lo tanto, un objeto cualquiera es una variable si y sólo si ese objeto es el signo $x$, o es el signo $x$ seguido de un número finito de palotes. Me parece que este es el modo natural de entender una definición así y por eso, a diferencia de otros autores, me doy por contento con la presente advertencia general y no la repito en cada caso.

## A. SINTAXIS

Las expresiones del CP1 están formadas por secuencias finitas o *filas* de signos tomados de una lista finita o *alfabeto* de ideogramas y letras.[15] Si la fila $\eta$ es una secuencia de $n$ signos ($n \geq 0$), decimos que $n$ es la *longitud* de $\eta$; simbólicamente, $\Lambda(\eta) = n$. Decimos que la fila $\eta$ es *más breve* que la fila $\zeta$, si $\Lambda(\eta) < \Lambda(\zeta)$. Si $\Lambda(\eta) = m$ y $\Lambda(\zeta) = n$, $\eta^\frown\zeta$ es la fila de longitud $m + n$ cuyo $h$-ésimo signo es el $h$-ésimo signo de $\eta$ ($1 \leq h \leq m$) y cuyo ($m$+$k$)-ésimo signo es el $k$-ésimo signo de $\zeta$ ($1 \leq k \leq n$). En vez de $\eta^\frown\zeta$ escribiré simplemente $\eta\zeta$.

*El alfabeto*. El alfabeto del CP1 consta de los siete ideogramas $\neg$, $\rightarrow$, $\forall$, (, ), $\shortmid$, *, y de las tres letras $x$, $a$, $P$. Los ideogramas se llaman, respectivamente, signo de *negación*, signo de *implicación* (o *flecha*), signo de *cuantificación universal*, paréntesis izquierdo, paréntesis derecho, palote y asterisco; pero, en aras de la brevedad, los usaré a ellos mismos para nombrarlos. Por ejemplo, si $\eta$ es una fila de longitud 3 y $\zeta$ es una fila de longitud 2, $(\eta_{\shortmid\shortmid} \rightarrow \zeta)$ es la fila de longitud 10 formada por un paréntesis izquierdo seguido del primer, segundo y tercer signo de $\eta$ (en ese orden), seguidos de dos palotes, seguidos de una flecha, seguida del primer y segundo signo de $\zeta$, seguidos de un paréntesis derecho. Los signos $\neg$ y $\rightarrow$ suelen llamarse *conectivos*.

[Obsérvese que, según las convenciones introducidas en los dos últimos párrafos una fila $\Phi$ de letras griegas entremezcladas o no con ideogramas del

[15] Algunos autores dicen 'cuerda' de signos en vez de 'fila', pero me parece esa una traducción demasiado literal y poco expresiva del término inglés 'string'.

cálculo representa una colección infinita de filas de signos del cálculo, especificada por la estructura de $\Phi$. Por ejemplo, la fila $(\alpha \to \forall\beta)$ representa todas las filas de signos del cálculo que empiezan con (, terminan con ) y contienen la fila $\to\forall$ entre dos filas cualesquiera (de longitud $\geq 0$). Cuando atribuyo propiedades sintácticas o semánticas a una tal fila $\Phi$, debe entenderse que la aseveración se refiere a cada fila representada por $\Phi$.]

*Vocabulario*. Las expresiones del CP1 combinan ideogramas con *palabras* formadas conforme a ciertas reglas precisas. Distinguimos dos clases principales de palabras del CP1.

(i) *Términos*: En el CP1 hay dos clases de términos, llamados *variables* y *constantes*.
La letra $x$ es una variable. Si $\xi$ es una variable, $\xi_{\mid}$ es una variable.
La letra $a$ es una constante. Si $\omega$ es una constante, $\omega_{\mid}$ es una constante.

(ii) *Predicados*: Distinguimos, para cada número natural $n \geq 0$, la clase de los *predicados n-ádicos*. (Suele llamárselos *variables proposicionales* si $n = 0$, *predicados monádicos* si $n = 1$ y *predicados poliádicos* si $n > 1$).
La letra $P$ es un predicado 0-ádico.
Si $\pi$ es un predicado $n$-ádico, $\pi_{\mid}$ es un *predicado n-ádico*.
Si $\pi$ es un prédicado $n$-ádico que no contiene palotes, $\pi^*$ es un predicado $(n+1)$-ádico.

Una *versión* del CP1 incluye siempre *todas* las variables y *todos* los predicados 0-ádicos.[16] Incluye además una parte de las constantes y una parte de los predicados $n$-ádicos para uno o más valores de $n > 0$. (Las partes en cuestión pueden ser propias o impropias y aun vacías). El conjunto de las constantes y predicados $n$-ádicos ($n > 0$) de una dada versión del CP1 constituye lo que llamaré el *léxico* de esa versión. Obsérvese que cada versión se

[16] Como se verá en la sección C. Tautologías, los predicados 0-ádicos sirven para representar fórmulas indeterminadas y, por ende, las proposiciones expresables con ellas. De ahí el nombre de *variables proposicionales* que, como dije, suele dárseles (con cierta impropiedad). Por eso también se llama *Cálculo proposicional* al fragmento del cálculo predicativo que puede construirse utilizando solamente predicados 0-ádicos, conectivos y paréntesis (estos son los únicos ingredientes de las fórmulas del CP1 que llamaré *verifuncionales*; véase su definición en la sección C).

distingue de las demás únicamente por su léxico.[17] Una versión del CP1 con léxico $\mathscr{L}_1$ es una *extensión* de otra con léxico $\mathscr{L}_2$ si y sólo si $\mathscr{L}_2 \subseteq \mathscr{L}_1$. La versión cuyo léxico contiene todas las constantes y todos los predicados *n*-ádicos ($n > 0$) del CP1 es la máxima extensión posible de cualquier otra versión. La llamo la *versión máxima* del CP1. Su léxico, $\mathscr{L}_{max}$, es el léxico máximo.

Sean η y ζ dos filas de signos (posiblemente de longitud 0) y sea θ la fila $\eta\frown\alpha\frown\zeta$, donde α es una variable, una constante o un predicado. Supongamos que ζ no comienza con un palote ni con un asterisco. Diremos entonces que la fila ζ determina una *posición* en θ y que α *figura* en θ en la *posición* determinada por ζ, o que α *ocupa* en θ la posición que precede a ζ.[18] Una misma variable, constante o predicado puede figurar en varias posiciones en una fila dada; estas posiciones se numeran —primera, segunda, tercera, . . . — según decrece la longitud de la fila que las determina.

*Fórmulas*. Una fila de signos formada por un predicado *n*-ádico seguido de *n* términos es una *fórmula elemental*. Toda fórmula elemental es una *fórmula*. Si α y β son fórmulas y ξ es una variable, también son *fórmulas* las filas de signos descritas a continuación:

| | |
|---|---|
| $[F_\neg]$ | $\neg\alpha$ |
| $[F_\rightarrow]$ | $(\alpha \rightarrow \beta)$ |
| $[F_\forall]$ | $\forall\xi\alpha$ |

La fila $\forall\xi$ es un *cuantificador* (*universal*) que *liga* la variable ξ. El *alcance* del cuantificador $\forall\xi$ en la fórmula $\forall\xi\alpha$ es precisamente la fórmula α que comienza inmediatamente a la derecha del mismo. Por aplicaciones repeti-

17 Chang y Keisler, en su espléndida *Model Theory* (tercera edición, 1990), llaman ‘lenguaje’ (*language*) a lo que yo llamo ‘léxico’. Ambas denominaciones se salen del uso común del respectivo vocablo, pero la adoptada por ellos me parece menos afortunada: si aprendo una palabra castellana que antes desconocía y al mismo tiempo me olvido de otra, no diríamos que he cambiado de lenguaje sino que ha variado mi léxico.

18 Exigimos que ζ no comience con palote ni asterisco para que no se pueda decir que la variable (constante, predicado) α figura en una cierta posición en θ cuando aparece allí formando parte de otra variable (constante, predicado). Conforme a nuestra estipulación, el predicado $P^{**}$ no figura en ninguna posición en la fórmula $P^{***}x_{||}x_{|}x_{|||}$ y la variable $x_|$ figura en una sola, a saber, la determinada por la fila $x_{|||}$.

das de la regla [$F_\forall$], puede obviamente ocurrir que una fórmula dada contenga más de un cuantificador que ligue a la misma variable. Sea $\beta$ una fórmula en que figura una variable $\xi$. Ésta puede ocupar tres tipos de posiciones en $\beta$: (A) inmediatamente a la derecha de un signo de cuantificación; (B) dentro del alcance de uno o más cuantificadores que ligan a $\xi$; (C) fuera de todo cuantificador y también fuera del alcance de todo cuantificador que ligue a $\xi$. Diremos que $\xi$ está *libre* en $\beta$ en cada posición del tipo C y que tal posición es una *posición libre* de $\xi$ en $\beta$. Diremos que $\xi$ está *ligada* por un determinado cuantificador que liga a $\xi$ (i) en la posición que ocupa dentro de él y (ii) en cada una de sus posiciones libres dentro del alcance de ese cuantificador.[19] Una fórmula que contiene una o más variables libres se dice *abierta*. Una fórmula que no contiene ninguna variable en una posición libre se dice *cerrada*. La *clausura universal* de una fórmula $\alpha$ que contiene las variables libres $\xi_1, \ldots, \xi_n$ (numeradas según el orden de su primera aparición en $\alpha$) es la fórmula cerrada $\forall\xi_1 \ldots \forall\xi_n\alpha$.

Si $\alpha$ y $\beta$ son fórmulas, decimos que $\beta$ es una *subfórmula* de $\alpha$ si $\alpha = \eta^\frown\beta^\frown\zeta$, donde $\eta$ y $\zeta$ son filas de signos (posiblemente de longitud 0) y $\zeta$ no empieza con un palote o un asterisco.[20] En tal caso, decimos también que $\beta$ *ocupa en* $\alpha$ *la posición determinada por* $\zeta$, o *que precede a* $\zeta$.

A veces hay que reemplazar en una fórmula dada una cierta variable libre por otro término (variable o constante). Nos conviene tener una expresión simbólica concisa que denote en nuestro idioma el producto de esta operación. Para evitar repeticiones, definiré tal expresión de modo que nos sirva no sólo mientras hablamos del CP1, en que todo término es idéntico a una constante o a una variable, sino también más tarde, al hablar de cálculos (como el CP1=) en que los términos pueden ser expresiones más complejas. Sea $\varphi$ una fórmula, $\xi$ una variable y $\tau$ un término del cálculo estudiado. Consideremos primero un caso sencillo: ninguna posición libre de $\xi$ en $\varphi$ está situada dentro del alcance de un cuantificador que ligue alguna de las

19 Por ejemplo, si la fórmula $\beta$ es la fila $\gamma\forall\xi\alpha\delta$ y $\alpha$ es el alcance del cuantificador $\forall\xi$ precedido inmediatamente por la fila $\gamma$, $\forall\xi$ sólo liga a $\xi$ dentro de sí y en las posiciones libres de $\xi$ en $\alpha$, pero no liga a $\xi$ en ninguna de las posiciones que $\xi$ ocupa en las filas $\gamma$ y $\delta$ ni en las posiciones de $\xi$ en $\alpha$ dentro de otro cuantificador que ligue a $\xi$ o dentro del alcance de tal cuantificador. (No se olvide que $\alpha$ podría ser una fila $\kappa\forall\xi\lambda$).

20 En virtud de la condición impuesta a $\zeta$, las fórmulas $P_{||}$ , $\neg P$ y $P{*}x$ *no son subfórmulas* de la formula $(P_{|||} \rightarrow \neg P{*}x_{||})$.

variables contenidas en $\tau$. Entonces, la expresión simbólica $\varphi_{\tau/\xi}$ designa (en castellano) la fórmula obtenida al insertar $\tau$ en vez de $\xi$ en todas las posiciones donde $\xi$ está libre en $\varphi$. Por cierto, si $\xi$ no figura en $\varphi$ o sólo figura en posiciones donde está ligada, $\varphi_{\tau/\xi} = \varphi$. En el caso general, algunas posiciones libres de $\xi$ en $\varphi$ se hallarán posiblemente dentro del alcance de cuantificadores que ligan a ciertas variables $\eta_1, \ldots, \eta_k$ que también figuran en el término $\tau$. Ordenemos las variables del cálculo según el número de palotes que contienen: $x$, $x_{|}$, $x_{||}$, $x_{|||}$, . . . . Sean $\zeta_1, \ldots, \zeta_k$ las $k$ primeras variables (en ese orden) que no figuran en la fórmula $\varphi$ ni en el término $\tau$. Sea $\varphi'$ la fórmula que se obtiene al reemplazar $\eta_i$ por $\zeta_i$ en todas las posiciones *ligadas* que $\eta_i$ ocupa en $\varphi$ ($1 \leq i \leq k$). Entonces, la expresión $\varphi_{\tau/\xi}$ designa la fórmula obtenida al insertar $\tau$ en vez de $\xi$ en todas las posiciones donde $\xi$ está libre en $\varphi'$.[21] En forma análoga puede introducirse la expresión $\varphi_{\tau_1,\ldots,\tau_\nu/\xi_1,\ldots,\xi_\nu}$ para simbolizar el producto de la sustitución *simultánea* del $\nu$-tuplo de variables $\langle\xi_1, \ldots ,\xi_\nu\rangle$ por el $\nu$-tuplo de términos $\langle\tau_1, \ldots ,\tau_\nu\rangle$ en todas las posiciones libres que dichas variables ocupan en $\varphi$. Suelo escribir $\phi(\tau/\xi)$ en vez de $\varphi_{\tau/\xi}$ y $\phi(\tau_1/\xi_1, \ldots ,\tau_\nu/\xi_\nu)$ en vez de $\varphi_{\tau_1,\ldots,\tau_\nu/\xi_1,\ldots,\xi_\nu}$.

## B. SEMÁNTICA

Como el lector habrá notado, las reglas sintácticas se refieren únicamente a signos y combinaciones de signos, esto es, a figuras inscritas en una pizarra o una hoja de papel (o, si se quiere, a los prototipos ideales de tales figuras), sin tener en cuenta lo que dichos signos significan. Las reglas semánticas se refieren al significado de las fórmulas construidas conforme a las reglas sintácticas. Específicamente, las reglas semánticas del CP1 establecen las condiciones necesarias y suficientes para que una dada fórmula de una versión particular de CP1 sea verdadera *relativamente a una cierta interpretación* de dicha versión. Sea, pues, CP1† una tal versión particular. Como ya

[21] En otras palabras: para transformar $\varphi$ en $\varphi_{\tau/\xi}$ primero reemplazo en $\varphi$ toda variable ligada que figure en $\tau$ por una variable *nueva*, la primera que no figure en $\varphi$ ni en $\tau$; así obtengo la fórmula $\varphi'$; luego reemplazo $\xi$ por $\tau$ en todas las posiciones libres de $\xi$ en $\varphi'$. Doy un ejemplo. Sea $\varphi$ la fórmula $\forall x(P^{**}x_{||}x \vee \exists x_{||}P^{***}xx_{||}x_{||||})$. Entonces, la primera variable que no figura en $\phi$ es $x_{|}$; aplicando la receta antedicha, tenemos pues que $\varphi_{x_{||}/x_{||||}}$ es la fórmula $\forall x(P^{**}x_{||}x \vee \exists x_{|}P^{***}xx_{|}x_{||})$.

he indicado, el vocabulario de CP1† incluye todos los ideogramas, variables y predicados 0-ádicos del CP1, más una selección bien definida de constantes y de predicados $n$-ádicos para ciertos valores de $n > 0$. Las reglas semánticas se formulan de modo que sean aplicables, como quiera que se haga esa selección.

Sea $\mathcal{D}$ una colección (no-vacía) de objetos cualesquiera. Sea $\mathcal{V}$ el conjunto de los términos y predicados del CP1† y sea $\mathcal{F}$ el conjunto de todas las fórmulas formadas con palabras de $\mathcal{V}$ conforme a nuestras reglas sintácticas. Una *interpretación* del CP1† es un par $\langle\mathcal{D},f\rangle$, donde $\mathcal{D}$ un conjunto (no-vacío) de objetos cualesquiera y $f$ es una aplicación, definida en $\mathcal{V} \cup \mathcal{F}$, que satisface las estipulaciones que se enuncian a continuación. ($\mathcal{D}$ se llama el *dominio* de la interpretación; también se suele decir que $f$ es una interpretación en el dominio $\mathcal{D}$).

[$I_\tau$] Para todo término $\tau \in \mathcal{V}$, $f(\tau) \in \mathcal{D}$.

[$I_0$] Para todo predicado 0-ádico $\pi \in \mathcal{V}$, $f(\pi) \in \{0,1\}$.

[$I_n$] Para todo predicado $n$-ádico $\pi \in \mathcal{V}$ $(n \geq 1)$, $f(\pi) \subseteq \mathcal{D}^n$ ($f(\pi)$ es un conjunto de $n$-tuplos de elementos de $\mathcal{D}$).

[$I_\alpha$] Si $\alpha$ es una fórmula formada por un predicado $n$-ádico $\pi \in \mathcal{V}$ seguido de los términos $\tau_1, \ldots, \tau_n$, $f(\alpha) = 0$ si $\langle f(\tau_1), \ldots, f(\tau_n)\rangle \in f(\pi)$ —esto es, si el $n$-tuplo formado por los objetos que $f$ asigna a los términos $\tau_1, \ldots, \tau_n$ es uno de los $n$-tuplos comprendidos en el conjunto que $f$ asigna al predicado $\pi$— y $f(\alpha) = 1$ si $\langle f(\tau_1), \ldots, f(\tau_n)\rangle \notin f(\pi)$.

[$I_\neg$] Si $\alpha$ es la fórmula $\neg\beta$, $f(\alpha) = 0$ si $f(\beta) = 1$ y $f(\alpha) = 1$ si $f(\beta) = 0$.

[$I_\rightarrow$] Si $\alpha$ es la fórmula $(\beta \rightarrow \gamma)$, $f(\alpha) = 0$ si $f(\beta) = 1$ ó $f(\gamma) = 0$, y $f(\alpha) = 1$ sólo si $f(\beta) = 0$ y $f(\gamma) = 1$.

[$I_\forall$] Sea $\alpha$ la fórmula $\forall\xi\beta$. Diré que $f'$ es una $\xi$-variante de $\langle\mathcal{D},f\rangle$ si $f'$ es una interpretación de CP1† en $\mathcal{D}$ y para cualquier predicado o término $\omega \neq \xi$ perteneciente al vocabulario de CP1†, $f'(\omega) = f(\omega)$. Entonces, $f(\alpha) = 0$ si para *toda* $f'$ que sea una $\xi$-variante de $f$, $f'(\beta) = 0$; y $f(\alpha) = 1$ si *hay una* $f'$, $\xi$-variante de $\langle\mathcal{D},f\rangle$, tal que $f'(\beta) = 1$.

Una fórmula $\alpha$ es *verdadera* en la interpretación $\langle\mathcal{D},f\rangle$ si y sólo si $f(\alpha) =$

0.[22] De otro modo, $\alpha$ es *falsa* en la interpretación $\langle\mathcal{D},f\rangle$. Si hay al menos una interpretación $\langle\mathcal{D},f\rangle$ en la cual la fórmula $\alpha$ es verdadera, decimos que $\alpha$ es una fórmula *realizable*. Si $\alpha$ es verdadera en *toda* interpretación, decimos que una fórmula $\alpha$ es *válida* (o *universalmente válida)*. Si $\alpha$ no es verdadera en *ninguna* interpretación decimos que $\alpha$ es *irrealizable*. Claramente, las fórmulas válidas representan lo que llamamos comúnmente *verdades lógicas*, aseveraciones verdaderas *en todo caso*, mientras que las fórmulas irrealizables representan aseveraciones *lógicamente falsas*, que bajo ninguna circunstancia pueden ser verdaderas. Por eso, decimos a veces 'fórmula lógicamente verdadera (falsa)' por 'fórmula válida (irrealizable)'. Decimos asimismo que una fórmula $\alpha$ es una *consecuencia lógica* o simplemente una *consecuencia* de un conjunto de fórmulas $\Delta$ (simbólicamente, $\Delta \vDash \alpha$), si $f(\alpha) = 0$ en toda interpretación $\langle\mathcal{D},f\rangle$ tal que $f(\beta) = 0$ para cada fórmula $\beta \in \Delta$. Si $\Delta \vDash \alpha$ y el conjunto $\Delta$ contiene una sola fórmula $\beta$, decimos también que $\alpha$ es una consecuencia de $\beta$ (simbólicamente, $\beta \vDash \alpha$). Si $\alpha$ es consecuencia de $\beta$ y viceversa, esto es, si en toda interpretación $\langle\mathcal{D},f\rangle$, $f(\alpha) = f(\beta)$, decimos que $\alpha$ y $\beta$ son *lógicamente equivalentes*. Obsérvese que, en virtud de nuestras definiciones, una fórmula $\alpha$ es válida si y sólo si $\varnothing \vDash \alpha$.

## C. TAUTOLOGÍAS

A principios de siglo solía usarse 'tautología' como sinónimo de 'verdad lógica' (o de 'fórmula válida'), pero aquí empleamos esa palabra en una acepción más restringida que procedo a explicar.

Como el lector habrá observado, una interpretación $\langle\mathcal{D},f\rangle$ asigna a cada predicado 0-ádico un valor 0 ó 1, pero no le atribuye propiamente un significado. Por otra parte, con arreglo a nuestras estipulaciones, cada predicado 0-ádico $\Pi$ es una fórmula verdadera en $\langle\mathcal{D},f\rangle$ si $f(\Pi) = 0$ y falsa en $f$ si $f(\Pi) = 1$. Asímismo, una fórmula $\alpha$ que sólo contenga predicados 0-ádicos se llama verdadera o falsa en $\langle\mathcal{D},f\rangle$ según el valor de $f(\alpha)$, calculable de

22 Obsérvese que, al revés de muchos libros de texto, asocio el número 0 a la verdad, el número 1 a su ausencia. Cualquier decisión en esta materia es, por cierto, arbitraria; la mía se inspira en palabras de la diosa que instruyó a Parménides, según las refiere Simplicio: ’Αληθείης εὐκυκλέος ἀτρεμὲς ἦτορ (Parménides, fr. 1.11 DK).

acuerdo con $[I_\neg]$ y $[I_\rightarrow]$. Aunque pueden parecer extravagantes y arbitrarias, estas prácticas tienen un sentido, como ahora veremos. También se entenderá por qué exigimos que cada versión del CP1 incluya *todos* los (infinitos) predicados 0-ádicos.

Digamos que una fórmula es una *fórmula verifuncional* si en ella no figuran variables, constantes, predicados monádicos o poliádicos ni cuantificadores. En otras palabras, una fórmula verifuncional sólo puede contener predicados 0-ádicos, conectivos y paréntesis. Cada predicado 0-ádico diferente que figura en una o más posiciones en una fórmula verifuncional $\alpha$ es un *componente* de $\alpha$. El epíteto cacofónico pero consagrado 'verifuncional' (del inglés *truth-functional*) se explica fácilmente. Sea $\langle \mathscr{D}, f \rangle$ una interpretación del CP1†. Si $\alpha$ es una fórmula verifuncional del CP1† con componentes $\Pi_1, \ldots, \Pi_n$, es claro que $f(\alpha)$ está unívocamente determinado por $f(\Pi_1), \ldots, f(\Pi_n)$, y en la jerga matemática esto suele expresarse diciendo que el *valor veritativo* (verdadero o falso) de $\alpha$ depende *funcionalmente* del valor veritativo de sus componentes.

Supongamos ahora que $\alpha$ es una fórmula cualquiera. Sea $\beta$ una fórmula verifuncional cuyos componentes $\Pi_1, \ldots, \Pi_n$ $(n \geq 1)$ no figuran en $\alpha$. Decimos que $\beta$ es un *prototipo verifuncional* de $\alpha$ si hay una lista de fórmulas diferentes $\langle \alpha_1, \ldots, \alpha_n \rangle$ tal que $\alpha$ es el resultado de reemplazar con $\alpha_k$ el componente $\Pi_k$ en todas las posiciones que $\Pi_k$ ocupa en $\beta$ $(1 \leq k \leq n)$. En tal caso diremos también que $\alpha$ se deriva de $\beta$ *por reemplazo* y que las fórmulas $\alpha_1, \ldots, \alpha_n$ son los *componentes verifuncionales* de $\alpha$ relativos al prototipo $\beta$. Obviamente el valor de $\alpha$ en una interpretación $\langle \mathscr{D}, f \rangle$ depende funcionalmente de $f(\alpha_1), \ldots, f(\alpha_n)$.

Ahora puedo explicar lo que en este libro se llama 'tautología'. Una fórmula $\alpha$ de uno de los cálculos predicativos descritos en este apéndice es una *tautología* si y sólo si $\alpha$ se deriva por reemplazo de una fórmula verifuncional válida $\beta$ (en cuyo caso, obviamente, $f(\alpha) = 0$ cualquiera que sea el valor de $f$ en los componentes verifuncionales de $\alpha$ relativos a $\beta$, y por ende también $\alpha$ es válida). El uso del término se extiende por analogía a los cálculos predicativos de que se habla en el cuerpo del libro.

## D. ACLARACIONES Y ABREVIATURAS

Se advertirá que una interpretación $\langle\mathcal{D},f\rangle$ de una versión cualquiera CP1† del CP1 asigna un objeto del dominio $\mathcal{D}$ a cada término del CP1†, el cual puede entonces considerarse como un nombre del objeto respectivo. Asigna asimismo una colección de objetos de $\mathcal{D}$ a cada predicado monádico, una colección de pares de objetos a cada predicado diádico, una colección de tríos a cada predicado triádico, etc., de suerte que cada predicado monádico del CP1† viene a designar la *propiedad* de pertenecer a la respectiva colección de objetos; y cada predicado poliádico, la *relación* que subsiste entre los componentes de un elemento de la respectiva colección de *n*-tuplos de objetos. Por último, la interpretación asigna a cada fórmula del CP1† uno de los valores 0 ó 1, que la última estipulación semántica asocia respectivamente a la verdad o falsedad de esa fórmula. De este modo, nuestras reglas semánticas caracterizan a las fórmulas del CP1† como oraciones declarativas.

Analicemos un ejemplo. Sean *a* una constante, $P^*$ un predicado monádico y $P^{**}$ un predicado diádico del CP1†. ¿En qué condiciones precisas es verdadera en $\langle\mathcal{D},f\rangle$ la fórmula $\forall x(P^*x \rightarrow P^{**}xa)$? De acuerdo con la regla $[I_\forall]$, $\forall x(P^*x \rightarrow P^{**}xa)$ es verdadera si y sólo si $(P^*x \rightarrow P^{**}xa)$ es verdadera en toda interpretación $\langle\mathcal{D},f'\rangle$ tal que $f'$ sea una *x*-variante de $\langle\mathcal{D},f\rangle$. Según la regla $[I_\rightarrow]$, $(P^*x \rightarrow P^{**}xa)$ es verdadera en $\langle\mathcal{D},f'\rangle$ a menos que $x \in f'(P^*)$ y $\langle x,a\rangle \notin f'(P^{**})$. Por lo tanto, para que la fórmula $\forall x(P^*x \rightarrow P^{**}xa)$ sea verdadera en la interpretación $\langle\mathcal{D},f\rangle$ es necesario y suficiente que cualquier objeto del dominio $\mathcal{D}$ que tenga la propiedad designada en esa interpretación por $P^*$ tenga con el objeto llamado *a* la relación designada por $P^{**}$.

He elegido para el CP1 un alfabeto paupérrimo, lo cual permite reducir la gramática a unas pocas reglas muy simples, pero dificulta la lectura. Para hacerla más fácil empleamos abreviaturas. Así, en vez de la fila de *r* palotes que individualizan un cierto término o predicado, escribimos el subíndice numérico *r* ($x_5$ en vez de $x_{|||||}$, etc.); en vez de la fila de *n* asteriscos que indican la *n*-adicidad de un predicado, escribimos el exponente numérico *n* ($P^4_3$ en vez de $P^{****}_{|||}$). Para mayor brevedad, podemos omitir los exponentes, ya que en una fórmula bien escrita la *n*-adicidad del predicado está indicada por el número de términos que lo acompañan. Asímismo, en contextos en que se usan pocas variables, constantes o predicados diferentes, escribimos *x*, *y*, *z*, *u*, *v*, *w* en vez de $x_0$, $x_1$, $x_2$, $x_3$, $x_4$, $x_5$; *a*, *b*, *c*, *d* en vez de

$a_0$, $a_1$, $a_2$, $a_3$; y cualquier mayúscula cursiva, sin subíndice ni exponente, en vez de *P* seguida de astericos y palotes.

Otras abreviaturas, combinadas con las reglas semánticas antedichas, engloban ciertas ideas lógicas corrientes. Por ejemplo, es claro que la fórmula $\neg\forall x\neg P^{**}xa$ es verdadera en la interpretación $\langle\mathcal{D},f\rangle$ si y sólo si hay por lo menos un objeto en $\mathcal{D}$ que tiene con el objeto llamado *a* la relación designada por $P^{**}$. Como esta condición es fácilmente comprensible por sí misma, se adopta la fila '$\exists\xi$' como abreviatura para representar a la fila '$\neg\forall\xi\neg$', que forman el signo de negación $\neg$, seguido por el signo de cuantificación universal $\forall$, seguido por una variable $\xi$, seguida por el signo de negación $\neg$. Se estipula asímismo que, si $\alpha$ y $\beta$ son fórmulas, la fórmula $(\neg\alpha \to \beta)$ se abrevia $(\alpha \vee \beta)$, la fórmula $\neg(\alpha \to \neg\beta)$ se abrevia $(\alpha \wedge \beta)$ y la fórmula $\neg((\alpha \to \beta) \to \neg(\alpha \to \beta))$ se abrevia $(\alpha \leftrightarrow \beta)$. Si el lector no está familiarizado con el tema, debe armarse ahora de papel y lápiz y verificar que las reglas semánticas del CP1 implican que:

[$I_\vee$] Si $\alpha$ es la fórmula $(\beta \vee \gamma)$, $f(\alpha) = 0$ si $f(\beta) = 0$ o $f(\gamma) = 0$, y $f(\alpha) = 1$ sólo si $f(\beta) = f(\gamma) = 1$.

[$I_\wedge$] Si $\alpha$ es la fórmula $(\beta \wedge \gamma)$, $f(\alpha) = 1$ si $f(\beta) = 1$ o $f(\gamma) = 1$, y $f(\alpha) = 0$ sólo si $f(\beta) = f(\gamma) = 0$.

[$I_\leftrightarrow$] Si $\alpha$ es la fórmula $(\beta \leftrightarrow \gamma)$, $f(\alpha) = 0$ si y sólo si $f(\beta) = f(\gamma)$, y $f(\alpha) = 1$ si y sólo si $f(\beta) \neq f(\gamma)$.

[$I_\exists$] Si $\alpha$ es la fórmula $\exists\xi\beta$, $f(\alpha) = 0$ si hay *alguna* $f'$, $\xi$-variante de $\langle\mathcal{D},f\rangle$, tal que $f'(\beta) = 0$; y $f(\alpha) = 1$ sólo si para *toda* $f'$ que sea una $\xi$-variante de $\langle\mathcal{D},f\rangle$, $f'(\beta) = 1$.

Los ideogramas $\vee$, $\wedge$, $\leftrightarrow$, son los signos (conectivos) de *disyunción*, *conjunción* y *equivalencia*, respectivamente; $\exists$ es el signo de *cuantificación existencial* ($\exists$ seguido de una variable es un *cuantificador existencial*).

Minimizar los ideogramas facilita el estudio de las propiedades generales de un cálculo, pues reduce el número de casos alternativos que hay que considerar para demostrarlas. Pero para el uso de un cálculo predicativo es más cómodo incluir también en su alfabeto "oficial" los cuatro ideogramas $\wedge$, $\vee$, $\leftrightarrow$ y $\exists$. Para ello hay que introducir reglas sintácticas [$F_\vee$], [$F_\wedge$], [$F_\leftrightarrow$] y [$F_\exists$] que legitimen las fórmulas $(\alpha \vee \beta)$, $(\alpha \wedge \beta)$, $(\alpha \leftrightarrow \beta)$ y $\exists\xi\alpha$, y agregar las reglas semánticas [$I_\vee$], [$I_\wedge$], [$I_\leftrightarrow$] y [$I_\exists$] a las estipulaciones que defi-

nen la aplicación $f$. *Tal es la convención que seguimos en el resto de este apéndice y, en general, en este libro*. Bajo dicha convención, por ejemplo, la expresión $(\alpha \vee \beta)$ no es ya una mera abreviatura de la fórmula $(\neg\alpha \rightarrow \beta)$ sino, al igual que ésta, una fórmula genuina del cálculo. Por otra parte, es claro que $(\alpha \vee \beta)$ y $(\neg\alpha \rightarrow \beta)$ son lógicamente equivalentes. El lector para quien esto no sea obvio debe ejercitarse en probarlo; debe probar también la equivalencia lógica de $(\alpha \wedge \beta)$ con $\neg(\alpha \rightarrow \neg\beta)$, de $(\alpha \leftrightarrow \beta)$ con $\neg((\alpha \rightarrow \beta) \rightarrow \neg(\alpha \rightarrow \beta))$ y de $\exists\xi\alpha$ con $\neg\forall\xi\neg\alpha$.

Ocasionalmente es útil emplear la negrita para abreviar una fila de símbolos consecutivos de la misma clase. Por ejemplo, en vez de $x_1x_2x_3x_4x_5$ escribimos simplemente $\mathbf{x}$ o, si deseamos mayor precisión, $\mathbf{x}^5$. Asímismo $\boldsymbol{\exists}\mathbf{x}^4$ y $\boldsymbol{\forall}\mathbf{x}^6$ representan, respectivamente, una fila de 4 cuantificadores existenciales y una de 6 cuantificadores universales.

## E. CÁLCULO PREDICATIVO CON IDENTIDAD

Una escritura conceptual sirve para expresar las matemáticas sólo si en ella se puede decir, llegado el caso, que tal objeto es el (único) valor de una cierta función para un dado argumento; por ejemplo, que 12 es la suma de 7 y 5. En el CP1 esto es posible si reservamos un predicado diádico para representar la relación de identidad, representamos mediante predicados $(n+1)$-ádicos las funciones —o, más generalmente, las aplicaciones— cuyos argumentos son $n$-tuplos, y acompañamos cada uso de estos predicados de una o más aseveraciones que los caractericen como tales. Pero este procedimiento no es cómodo. Así, por ejemplo, si abreviamos con $I$ el predicado diádico con que representaremos la identidad, tenemos que postular, *para cada fórmula* $\alpha$, la aseveración $\forall x\forall y(Ixy \rightarrow (\alpha \leftrightarrow \alpha_{x/y}))$. Si $F$ es un predicado triádico elegido para representar una aplicación cuyo dominio es un conjunto de pares, hay que incluir la condición $\forall x\forall y\forall z(Fxyz \rightarrow \forall w(Fxyw \rightarrow Izw))$. Más práctico y, en cierto modo, más natural, es agregar al alfabeto del CP1 un ideograma que signifique la identidad y un método para construir términos que denoten el valor de aplicaciones en cada argumento dado y postular nuevas reglas sintácticas y semánticas que gobiernen el uso de estos signos. Llamaré *cálculo predicativo con identidad* o CP1= a la escritura conceptual así ampliada. He aquí la gramática del CP1=:

*El alfabeto*. El alfabeto del CP1= consta de los doce ideogramas $\neg$, $\rightarrow$, $\wedge$, $\vee$, $\leftrightarrow$, $\forall$, $\exists$, =, (, ), $_{\text{|}}$, *, y de las cuatro letras *x*, *a*, *f*, *P*. El ideograma = es el *signo de identidad*.

*Vocabulario*.

(i) *Functores*: Distinguimos, para cada entero positivo $n \geq 1$, la clase de los *functores n-arios*. *f** —la letra *f* seguida de un asterisco— es un functor 1-ario. Si $\varphi$ es un functor *n*-ario ($n \geq 1$), $\varphi_{|}$ es un functor *n*-ario. Si $\varphi$ es un functor *n*-ario que no contiene palotes, $\varphi$* es un functor (*n*+1)-ario.

(ii) *Términos*: Las *variables* y *constantes* del CP1= se definen como las del CP1. Además, cualquier functor *n*-ario seguido de *n* términos es un término.

(iii) *Predicados*: Se definen como en el CP1.

*Léxicos*. Un conjunto (vacío, finito o infinito) de constantes, predicados *n*-ádicos y functores *m*-ádicos (para uno o más valores positivos de *n* y *m*) constituye el *léxico* de una *versión* del CP1=.

*Fórmulas*. A las reglas de formación de fórmulas propias del CP1 hay que agregar la siguiente: Si $\tau$ y $\sigma$ son términos, la fila de signos descrita a continuación es una fórmula simple:

$$[F_{=}] \qquad (\tau = \sigma)$$

*Semántica*. Sea CP1=† una versión del cálculo predicativo de primer orden con identidad. Sea $\langle\mathcal{D},f\rangle$ una interpretación de CP1=†. Sea $\mathcal{V}$ el conjunto de los functores, variables, constantes y predicados de CP1=†. Se mantienen todas las reglas semánticas que dimos para CP1†, excepto $[I_{\tau}]$ que se reemplaza por la regla $[I'_{\tau}]$ enunciada abajo. Además, se agregan las reglas $[I_0]$ y $[I_{=}]$. Las nuevas reglas determinan la interpretación de los términos precedidos por functores y de las fórmulas construidas según la regla $[F_{=}]$.

$[I_0]$ Para todo functor *n*-ario $\varphi \in \mathcal{V}$, $f(\varphi)$ es una aplicación cuyo dominio es el producto cartesiano de *n* partes de $\mathcal{D}$ —no necesariamente distintas— no vacías y bien definidas, y cuyo codominio es una parte

bien definida de $\mathcal{D}$.

[I′$_\tau$] Si $\tau$ es una variable o una constante, $f(\tau) \in \mathcal{D}$. Si $\tau$ es un término formado por un functor $n$-ario $\varphi$ seguido de $n$ términos $\sigma_1, \ldots, \sigma_n$, entonces (i) si $\langle f(\sigma_1), \ldots, f(\sigma_n)\rangle$ pertenece al dominio de $f(\varphi)$, $f(\tau)$ es el valor de la aplicación $f(\varphi)$ en el argumento $\langle f(\sigma_1), \ldots, f(\sigma_n)\rangle$; y (ii) si $\langle f(\sigma_1), \ldots, f(\sigma_n)\rangle$ no pertenece al dominio de $f(\varphi)$, $f(\tau)$ no está definida.

[I$_=$] Si $\alpha$ es la fórmula $(\tau = \sigma)$, $f(\alpha) = 0$ si $f(\tau) = f(\sigma)$ y $f(\alpha) = 1$ si $f(\tau) \neq f(\sigma)$.[23]

## F. DECISIÓN Y DEDUCCIÓN

El cálculo predicativo y otros sistemas similares de “lógica simbólica” se inventaron ante todo para poder juzgar si un razonamiento es correcto o incorrecto por la sola inspección visual de los signos que lo representan (cf. la cita de Leibniz en la p. 129). Evidentemente, este propósito se habrá logrado con todos los razonamientos representables en el CP1= si se dispone de un algoritmo que permita calcular si una fórmula dada $\alpha$ es o no es válida. En efecto, si las premisas de un razonamiento se pueden representar con las fórmulas $\pi_1, \pi_2, \ldots, \pi_n$ y su conclusión con la fórmula $\kappa$, el razonamiento en cuestión será correcto si y sólo si la fórmula $(\pi_1 \wedge \pi_2 \wedge \ldots \wedge \pi_n) \rightarrow \kappa$ es válida. (El lector para quien esto no sea enteramente obvio debe, como siempre, comprobarlo con papel y lápiz: muestre que, si la forma antedicha no es válida, puede ocurrir que la fórmula $\kappa$ sea falsa aunque todas las fórmulas $\pi_1, \pi_2, \ldots, \pi_n$ sean verdaderas).

23 El lector alerta advertirá que en el enunciado de la regla [I$_=$] el signo ‘=’ que figura entre $\tau$ y $\sigma$ no es igual al signo ‘=’ que figura tres veces después de la coma. El signo más largo es el signo de igualdad del **CP1**=; el signo más breve es el familiar signo de igualdad común al castellano y los demás idiomas modernos. Este distingo no se observa en la mayor parte de este libro, en que el signo corriente ‘=’ se emplea como signo de igualdad en cualquier contexto, sea formal o informal. Con todo, uso el signo de igualdad largo cuando me refiero a la versión final de la escritura conceptual BS de Frege (Capitulo 2.2, Sección 2.3.3 y Apéndice XII), en la cual dicho signo representa a la vez la identidad entre objetos y la equivalencia entre oraciones (la cual presupone la identidad de sus valores veritativos).

Un algoritmo para determinar la validez de las fórmulas de un cálculo lógico (o de un fragmento de un cálculo lógico) se llama un *procedimiento de decisión* para ese cálculo (o fragmento). Para las fórmulas verifuncionales del CP1= hay un procedimiento de decisión —propuesto ya en 1880 por Peirce (CP, 3.387) y difundido en este siglo por Post (1921) y Wittgenstein (1922)— que consiste en la construcción de lo que se llaman "tablas de verdad". Dicho procedimiento se funda en los hechos siguientes: (i) Cada una de las reglas semánticas $[I_{\neg}]$, $[I_{\rightarrow}]$, $[I_{\vee}]$, $[I_{\wedge}]$ y $[I_{\leftrightarrow}]$ puede leerse como la definición de una operación sobre el conjunto {0,1} y permite calcular simplemente el valor asignado por esa operación a cada argumento admisible (que son dos en el caso de la operación 1-aria definida por $[I_{\neg}]$, cuatro en cada uno de los casos restantes).[24] (ii) Si α es una fórmula verifuncional con *n* componentes distintos $\alpha_1, \ldots, \alpha_n$, cada interpretación *f* pertenece a una y sólo una de $2^n$ clases de interpretaciones posibles, determinada por la lista de valores $f(\alpha_1), \ldots, f(\alpha_n)$. Computando $2^n$ veces $f(\alpha)$ con arreglo a lo dicho bajo (i), se sabrá si α es válida (esto es, si $f(\alpha) = 0$ en todos los casos).[25] Obviamente si α es válida y es un prototipo verifuncional de una fórmula β, también β es válida.

En los años 20, varios seguidores de Hilbert buscaron afanosamente un procedimiento de decisión para el CP1. Mas, como se explica en el Capítulo 2.11, no puede haber tal procedimiento (al menos, si nos atenemos a la definición precisa de 'algoritmo' propuesta por Church e insuperada hasta ahora). Hay, con todo, otro camino —más incierto, pero más viable— para verificar si un razonamiento es correcto o no, por la sola inspección visual de su representación simbólica. Este camino, que ya fue seguido por Frege (Capítulo 2.2), consiste en organizar el cálculo lógico como un *sistema deductivo*, premunido de axiomas y reglas de inferencia. Un *axioma* es una fórmula distinguida del cálculo. El conjunto de los axiomas puede ser finito

24 Simbolizando cada operación con el conectivo correspondiente, escribimos $\neg 0 = 1$, $\neg 1 = 0$; $0 \rightarrow 0 = 0$, $0 \rightarrow 1 = 1$, $1 \rightarrow 0 = 0$, $1 \rightarrow 1 = 0$; etc.

25 Supongamos que α tiene la estructura $(\alpha_1 \rightarrow (\alpha_2 \rightarrow \alpha_1))$. Para saber si α es o no válida hay que calcular $f(\alpha)$ para los cuatro valores posibles del par $\langle f(\alpha_1), f(\alpha_2)\rangle$, a saber, $\langle 0,0\rangle$, $\langle 0,1\rangle$, $\langle 1,0\rangle$ y $\langle 1,1\rangle$. Reemplazando $\alpha_1$ y $\alpha_2$ en α por el valor correspondiente y utilizando la "tabla de multiplicar para el operador →" presentada en la nota 23, calculamos que $f(\alpha) = 0$ en todos los casos. Por lo tanto, cualquier fórmula de la estructura indicada es una tautología.

o infinito, pero en todo caso tiene que haber un procedimiento que permita decidir, por la sola inspección visual de una fórmula, si ella es o no es un axioma. Por ejemplo, podemos estipular que cualquier fórmula obtenida reemplazando $\alpha$ por una fórmula del cálculo en la fila de signos $(\alpha \vee \neg\alpha)$ es un axioma. Una *regla de inferencia n–aria* ($n \geq 1$) es una receta para asociar una fórmula llamada *conclusión* a un $n$-tuplo de fórmulas llamadas *premisas*, atendiendo exclusivamente al aspecto visual de éstas y aquélla. Esto quiere decir que cada regla $n$–aria $\mathcal{R}$ provee un algoritmo para decidir si una fórmula dada es o no una conclusión, según $\mathcal{R}$, de un cierto $n$-tuplo de premisas. Por ejemplo, la regla de inferencia por *modus ponens* prescribe que, si $\alpha$ y $\beta$ son dos fórmulas cualesquiera, $\beta$ se puede asociar como conclusión al par de premisas $\langle\alpha, (\alpha \rightarrow \beta)\rangle$;[26] y obviamente, basta *deletrear* un trío cualquiera de fórmulas para constatar si una de ellas es o no la conclusión por *modus ponens* de las otras dos.

Una vez que el cálculo se ha organizado como sistema deductivo, se puede definir lo que llamaremos una deducción o prueba. Sea $\mathcal{H}$ un conjunto de fórmulas. Una lista finita de fórmulas $\alpha_1, \ldots, \alpha_n$ constituye una *prueba* de la fórmula $\alpha_n$ *a partir de las hipótesis* $\mathcal{H}$ si, para cada $k$ ($1 \leq k \leq n$), o bien $\alpha_k \in \mathcal{H}$, o bien $\alpha_k$ es un axioma, o bien hay una regla de inferencia $r$-aria ($r < k$) que asocia $\alpha_k$ como conclusión a un $r$-tuplo de premisas contenidas en la lista $\alpha_1, \ldots, \alpha_{k-1}$. Si $\mathcal{H} = \varnothing$, decimos simplemente que $\alpha_1, \ldots, \alpha_n$ es una *prueba* de $\alpha_n$. A la luz de esta definición es claro que, dada la índole de los axiomas y las reglas de inferencia, hay un algoritmo para decidir si la lista $\alpha_1, \ldots, \alpha_n$ constituye o no una *prueba* de $\alpha_n$ a partir de determinadas hipótesis. Si hay una prueba de la fórmula $\alpha$ a partir de las hipótesis $\mathcal{H}$ decimos que, en el sistema deductivo adoptado, $\alpha$ *se deduce de* $\mathcal{H}$ o *es deducible de* $\mathcal{H}$; simbólicamente: $\mathcal{H} \vdash \alpha$. Si los axiomas se eligen entre las fórmulas de validez conocida, y las reglas de inferencia se diseñan de modo que la conclusión sea siempre una consecuencia lógica de las premisas, diremos que el sistema deductivo es *correcto*. En tal caso, como es obvio, $\alpha$ se deduce de $\mathcal{H}$ sólo si $\alpha$ es una consecuencia lógica de $\mathcal{H}$ ($\mathcal{H} \vdash \alpha$ sólo si

26 Recuérdese que según la convención adoptada '$(\alpha \rightarrow \beta)$' designa la fila de signos que forman el paréntesis izquierdo, seguido por la fila $\alpha$, seguida por la flecha, seguida por la fila $\beta$, seguida por el paréntesis derecho.

$\mathcal{H} \vDash \alpha$). En particular, sólo si $\alpha$ es válida hay una prueba de $\alpha$ (sin hipótesis).

Un sistema deductivo incorrecto sería, por cierto, inservible para el propósito original del cálculo lógico. Pero un sistema correcto sólo acabará de satisfacernos si a la vez es *completo*, esto es, si toda fórmula $\alpha$ que sea una consecuencia lógica de un conjunto de fórmulas $\mathcal{H}$ se deduce de $\mathcal{H}$ en el sistema. Entonces, $\mathcal{H} \vdash \alpha$ si y sólo si $\mathcal{H} \vDash \alpha$, y la relación puramente sintáctica de deducibilidad simula cabalmente la relación semántica de consecuencia lógica. Gödel (1930) demostró que el sistema deductivo ostensiblemente correcto para el CP1= que él toma de Russell y Whitehead es un sistema deductivo completo (Capítulo 2.8).

## G. UN SISTEMA DEDUCTIVO PARA EL CÁLCULO DE PRIMER ORDEN

A modo de ejemplo, propongo a continuación un sistema deductivo para el CP1, basado en uno de los sistemas de "deducción natural" introducidos por Gentzen en su tesis doctoral, "Investigaciones sobre la deducción lógica" (1934). Para mayor claridad, incluyo en el alfabeto —además de los signos indicados en la sección A— los conectivos $\wedge$ y $\vee$ y el signo de cuantificación $\exists$ (el significado de estos signos se explicó en la sección D). La característica más notable de los cálculos de Gentzen es que los componentes de una deducción —premisas, conclusión, etapas intermedias— no representan, por regla general, aseveraciones incondicionales, sino aseveraciones condicionadas por ciertos supuestos o hipótesis, tal como ocurre en la vida real y especialmente en el razonamiento matemático.[27] Por eso, tales componentes no son fórmulas, sino filas de fórmulas estructuradas en cierta forma. Para indicar esa estructura, agregamos al CP1 dos signos de puntuación: la coma [,] y la flecha secuencial [➛]. Un *secuente* es una fila de fórmulas del

[27] Como bien observa Gentzen, "una prueba matemática no está, en general, construida de un modo tan simple, que en ella se avance, mediante inferencias, de aseveraciones *válidas* a nuevas aseveraciones *válidas*. Ocurre también que una aseveración se *supone* válida y de ella se *infieren* nuevas aseveraciones cuya validez *depende* entonces de la validez de esa suposición" (1935, p. 511).

CP1 separadas por comas, seguida del signo ➛, seguida de una fórmula del CP1.[28] Las fórmulas que preceden al signo ➛ son las *prefórmulas* del secuente; la fórmula precedida por el signo ➛ es la *posfórmula*. La semántica de los secuentes es muy simple. Consideremos una determinada versión del CP1. Sea φ una fórmula de esa versión y Γ una fila de fórmulas separadas por comas. El secuente Γ ➛ φ es válido si y sólo si φ es verdadero en cada interpretación (de la versión considerada) en que ninguna de las fórmulas de Γ sea falsa. En virtud de esto, las prefórmulas de un secuente válido representan hipótesis cuya verdad en una interpretación dada asegura la verdad —en esa misma interpretación— de la aseveración representada por la posfórmula.

Llamo *secuente básico* a cualquier secuente de la forma φ ➛ φ, donde φ es una fórmula cualquiera. Cada secuente básico es válido de un modo trivial y puede usarse como axioma en las deducciones. No se admiten otros axiomas *lógicos*. El poder deductivo del sistema radica, pues, enteramente en sus reglas de inferencia. Cada regla de inferencia determina una relación entre una, dos o tres premisas y una conclusión. Llamo *inferencia* al conjunto ordenado que forman premisa(s) y conclusión. Distinguimos (i) *reglas estructurales* (que llevan de una premisa con cierta estructura secuencial a una conclusión de estructura diferente), (ii) *reglas de eliminación e introducción* de los conectivos binarios y los signos de cuantificación y (iii) dos reglas peculiares a la negación. En la siguiente formulación esquemática de las reglas de inferencia, las letras griegas mayúsculas Γ y Δ representan filas —posiblemente vacías— de fórmulas separadas por comas, las minúsculas φ, χ, ψ representan fórmulas, la minúscula τ representa un término, las minúsculas ξ y η representan variables, y los signos lógicos y de puntuación están, como de costumbre, representados por ellos mismos. Cada regla se presenta mediante una figura con una raya horizontal en el centro; el esquema de la o las premisas va sobre la raya, el esquema de la conclusión bajo la raya. Las inferencias por eliminación de ∧ e introducción de ∨ tienen dos figuras cada una.

[28] Gentzen dice ‘Sequenz’, que es la forma alemana de la misma palabra latina que en nuestro idioma se ha convertido en ‘secuencia’. Pero lo que nosotros llamamos ‘secuencia’ se llama ‘Folge’ en alemán. Por eso digo ‘secuente’ por ‘Sequenz’, siguiendo el ejemplo de los escritores de habla inglesa, que dicen ‘sequent’ (no ‘sequence’).

*Reglas estructurales*:

DEBILITAMIENTO

$$\frac{\Gamma \Rightarrow \chi}{\varphi, \Gamma \Rightarrow \chi}$$

CONTRACCIÓN

$$\frac{\varphi, \varphi, \Gamma \Rightarrow \chi}{\varphi, \Gamma \Rightarrow \chi}$$

PERMUTACIÓN

$$\frac{\Gamma, \varphi, \psi, \Delta \Rightarrow \chi}{\Gamma, \psi, \varphi, \Delta \Rightarrow \chi}$$

*Reglas de introducción y eliminación*:

| | INTRODUCCIÓN | ELIMINACIÓN |
|---|---|---|
| $\wedge$ | $\dfrac{\Gamma \Rightarrow \varphi \qquad \Gamma \Rightarrow \psi}{\Gamma \Rightarrow \varphi \wedge \psi}$ | $\dfrac{\Gamma \Rightarrow \varphi \wedge \psi}{\Gamma \Rightarrow \varphi}$ $\quad$ $\dfrac{\Gamma \Rightarrow \varphi \wedge \psi}{\Gamma \Rightarrow \psi}$ |
| $\vee$ | $\dfrac{\Gamma \Rightarrow \varphi}{\Gamma \Rightarrow \varphi \vee \psi}$ $\quad$ $\dfrac{\Gamma \Rightarrow \psi}{\Gamma \Rightarrow \varphi \vee \psi}$ | $\dfrac{\Gamma \Rightarrow \varphi \vee \psi \quad \varphi, \Delta \Rightarrow \chi \quad \psi, \Delta \Rightarrow \chi}{\Gamma, \Delta \Rightarrow \chi}$ |
| $\rightarrow$ | $\dfrac{\varphi, \Gamma \Rightarrow \psi}{\Gamma \Rightarrow (\varphi \rightarrow \psi)}$ | $\dfrac{\Gamma \Rightarrow \varphi \qquad \Delta \Rightarrow (\varphi \rightarrow \psi)}{\Gamma, \Delta \Rightarrow \psi}$ |
| $\forall$ | $\dfrac{\Gamma \Rightarrow \varphi_{\eta/\xi}}{\Gamma \Rightarrow \forall\xi\varphi}$ siempre que $\eta$ no figure en $\Gamma$ ni en $\forall\xi\varphi$ | $\dfrac{\Gamma \Rightarrow \forall\xi\varphi}{\Gamma \Rightarrow \varphi_{\tau/\xi}}$ donde $\tau$ es un término cualquiera |
| $\exists$ | $\dfrac{\Gamma \Rightarrow \varphi_{\tau/\xi}}{\Gamma \Rightarrow \exists\xi\varphi}$ donde $\tau$ es un término cualquiera | $\dfrac{\Gamma \Rightarrow \exists\xi\varphi \qquad \varphi_{\eta/\xi}, \Delta \Rightarrow \psi}{\Gamma, \Delta \Rightarrow \psi}$ siempre que $\eta$ no figure en $\Gamma$, $\Delta$, $\psi$ ni $\exists\xi\varphi$ |

*Reglas de la negación*:

REFUTACIÓN

$$\frac{\varphi, \Gamma \Rightarrow \psi \qquad \varphi, \Delta \Rightarrow \neg\psi}{\Gamma, \Delta \Rightarrow \neg\varphi}$$

ELIMINACIÓN DE $\neg\neg$

$$\frac{\Gamma \Rightarrow \neg\neg\varphi}{\Gamma \Rightarrow \psi}$$

Con un poco de reflexión el lector podrá comprobar que, si las premisas de cada una de las inferencias arriba representadas son secuentes válidos, la conclusión respectiva también es un secuente válido. Si no conocía el tema de antemano, debe poner especial empeño en entender por qué la restricción impuesta a la variable $\eta$ en las reglas de introducción de $\forall$ y de eliminación de $\exists$ (se exige que $\eta$ sea una variable "nueva") es a la vez necesaria y suficiente para que las inferencias gobernadas por esas reglas sean correctas.

Una prueba (Gentzen dice *Herleitung*, 'derivación') es una lista finita de secuentes cada uno de los cuales es un secuente básico o es la conclusión de una inferencia (conforme a alguna de las reglas indicadas) cuya(s) premisa(s) lo precede(n) en la lista. En otras palabras, si $\Delta = \langle S_0, \ldots, S_n \rangle$ es una lista finita de secuentes ($n \geq 0$), $\Delta$ es una *prueba* si y sólo si, para cada índice $k$ ($0 \leq k \leq n$), tenemos que $S_k$ es un secuente básico o la conclusión de una inferencia con premisas $S_h$, $S_i$, $S_j$ ($0 \leq h \leq i \leq j < k$).

Ilustraré estas ideas con un ejemplo modesto, una prueba de la conocida tautología $(P \rightarrow (P_1 \rightarrow P))$. La prueba va en la columna del centro; los números de la izquierda sirven para identificar a los secuentes; la columna de la derecha indica la regla en que se basa la introducción del secuente respectivo, así como el número de identidad de la premisa correspondiente. SB significa "secuente básico".

| | | |
|---|---|---|
| (1) | $P \succ P$ | SB |
| (2) | $P_1, P \succ P$ | Debilit., (1) |
| (3) | $P \succ (P_1 \rightarrow P)$ | Int. $\rightarrow$, (2) |
| (4) | $\succ (P \rightarrow (P_1 \rightarrow P))$ | Int. $\rightarrow$, (3) |

El secuente (1) es válido. La prueba se ajusta a las reglas. Por lo tanto, el secuente (4) es válido. Su posfórmula es, por ende, verdadera en cualquier interpretación del cálculo en que no sea falsa una de sus prefórmulas. Como el secuente (4) no tiene prefórmulas, no hay ninguna interpretación en que una de ellas sea falsa. Por lo tanto, la posfórmula $(P \rightarrow (P_1 \rightarrow P))$ es válida (verdadera en cualquier interpretación).

## H. CÁLCULO PREDICATIVO DE SEGUNDO ORDEN

Como se dijo, un cálculo predicativo de segundo orden debe permitirnos expresar generalizaciones acerca de propiedades y relaciones de los objetos individuales a que se refiere nuestro discurso. Ello demanda muy pocas innovaciones sintácticas.

*Alfabeto*. Además de los ocho ideogramas y cuatro letras del CP1=, incluimos las letras *X* y *F*.

*Variables*. Distinguimos tres clases de variables: variables *individuales*, formadas como en el CP1= por la letra *x* seguida o no de palotes; variables *funcionales* y variables *predicativas*. Los nuevos tipos de variables se caracterizan así: Una variable predicativa *n*-ádica consiste de la letra *X*, seguida de *n* asteriscos ($n \geq 1$), seguidos por 0 o más palotes. Una variable funcional *r*-aria consiste de la letra *F* seguida de *r* asteriscos ($r \geq 1$), seguidos por 0 o más palotes.

Esta innovación justifica los siguientes cambios en nuestra terminología: Llamamos *constantes* a las antiguas constantes, así como a los antiguos predicados y functores. Las constantes que empiezan con la letra *a* son *constantes individuales*, las que empiezan con la letra *f* son *constantes funcionales* y las que empiezan con la letra *P* son *constantes predicativas*. Llamamos *signos individuales* a las variables y constantes individuales, *functores* a las variables y constantes funcionales y *predicados* a las variables y constantes predicativas.

*Functores*. Conforme a la estipulación anterior, la segunda cláusula de la regla (i) concerniente al vocabulario del CP1= debe reemplazarse por ésta: "*f** y *F** son functores 1-arios". El resto queda igual.

*Términos*. Se mantiene lo dicho arriba, entendiéndose ahora que las variables y constantes aludidas bajo este título en la sintaxis del CP1 son variables y constantes *individuales*, y que en el párrafo homónimo de la sintaxis del CP1=, 'functor' significa 'variable o constante funcional'.

*Predicados*. Para que todas las variables predicativas queden comprendidas en esta categoría basta insertar esta cláusula en la regla pertinente del CP1: "La letra *X* es un predicado 0-ádico".

*Fórmulas*. En las reglas [$F_\forall$] y [$F_\exists$] se entiende ahora que $\xi$ designa una variable de cualquier clase, individual, funcional o predicativa.

Llamaré CP2= a la familia de cálculos de segundo orden con identidad gobernado por estas reglas de sintaxis (además a las ya formuladas para el CP1 y el CP1=). Una versión del CP2= incluye *todas* las variables de las tres clases, y un cierto subconjunto —posiblemente finito o incluso vacío— de las constantes, los functores y los predicados.

La semántica de los cálculos de segundo orden no se deja precisar con tanta facilidad como la sintaxis. Además de la llamada semántica estándar, descrita enseguida, se han propuesto otras muy diferentes (Henkin 1950).

Una interpretación estándar del CP2= se obtiene extendiendo a las nuevas variables (funcionales y predicativas) la función $f$ característica de una interpretación del CP1=. Específicamente, una interpretación $\langle \mathcal{D}, f \rangle$ asigna un objeto del dominio no vacío $\mathcal{D}$ a cada signo individual, un conjunto de $n$-tuplos de objetos de $\mathcal{D}$ a cada predicado $n$-ádico (constante o variable) y una aplicación de $\mathcal{D}^n$ en $\mathcal{D}$ a cada functor $n$-ario (constante o variable). Las reglas de semántica del CP1= se pueden mantener inalteradas, con tal que las palabras 'functor', 'término', 'predicado' y 'variable' se entiendan en sus nuevas acepciones.

## I. ALCANCES

En la Sección B definí el *alcance* de un cuantificador. Este concepto de alcance puede extenderse también a los conectivos, predicados y functores. En las definiciones siguientes me refiero en cada caso a *una* ocasión en que el objeto de que se habla figura en una fórmula dada. El alcance de un predicado $n$-ádico (o de un functor $n$-ario) es la fila de $n$ términos precedida inmediatamente por él. El alcance de un cuantificador o de un signo de negación es la subfórmula que comienza inmediatamente a su derecha. El alcance de un conectivo que no sea el signo de negación son las dos subfórmulas que lo flanquean. El lector que tenga dudas sobre estas definiciones debe hacer ejercicios. Diga, por ejemplo, cuáles son, respectivamente, los alcances de $\forall x_1$, $\exists x_2$, $\neg$, $\vee$, $P^2$, $P^3$ y $f^2$ en la fórmula

$$\neg \forall x_1 (P^3 x_1 x_2 x_3 \vee \exists x_2 \neg P^2 x_2 f^2 x_1 x_3).$$[29]

[29] Solución en la p. 540.

## J. OMISIÓN DE PARÉNTESIS

Para alivianar la lectura, en general no se imprimen todos los paréntesis que, según la sintaxis, pertenecen a una fórmula dada. Desde luego, se omite el primer y el último paréntesis de una fórmula cuando son, respectivamente, su primer y su último símbolo: en vez de $(\neg\alpha \to \beta)$, escribimos $\neg\alpha \to \beta$. Estipulamos un orden de prioridad de las asociaciones que los conectivos establecen entre las expresiones inmediatamente adyacentes. La conjunción $\wedge$ y la disyunción $\vee$ asocian las fórmulas adyacentes más estrechamente que la flecha $\to$ y la doble flecha $\leftrightarrow$. Según esto, $\alpha \wedge \beta \to \beta \vee \gamma$ es la fórmula $((\alpha \wedge \beta) \to (\beta \vee \gamma))$. Si una fórmula figura entre dos flechas, la asociación determinada por la flecha a su derecha tiene prioridad sobre la asociación determinada por la flecha a su izquierda. Así, $\alpha \to \beta \to \gamma \to \delta$ es la fórmula $(\alpha \to (\beta \to (\gamma \to \delta)))$. Por último, es corriente omitir los paréntesis que, según la regla $[F_=]$, rodean a las expresiones unidas por el signo de identidad $=$. No hay riesgo de confusión, porque, a diferencia de los conectivos $\wedge$, $\vee$, $\to$ y $\leftrightarrow$, la identidad se intercala entre *términos*, no entre *fórmulas*. Por lo tanto, si $\sigma$ y $\tau$ son términos y $\alpha$ y $\beta$ son fórmulas, en vez de $((\sigma = \tau) \to (\alpha \to \beta))$, podemos escribir sin ambigüedad $\sigma = \tau \to \alpha \to \beta$. La convención relativa al signo $=$ también nos permite omitir los paréntesis en $\neg(\sigma = \tau)$ y escribir $\neg\sigma = \tau$; pero lo corriente en este caso es usar el signo $\neq$ y escribir simplemente $\sigma \neq \tau$.

# X

# AXIOMAS DE LA LÓGICA
# (FREGE 1879)

Las nueve fórmulas de la escritura conceptual BS reproducidas a continuación son las leyes lógicas que Frege (1879) adopta como premisas indemostradas de sus deducciones. Están numeradas como en el libro de Frege. El lector hará bien en verificar que equivalen, respectivamente, a las nueve fórmulas del CP1 señaladas con los mismos números hacia el final del Capítulo 2.2. También debe verificar que esas nueve fórmulas son verdades lógicas conforme a las reglas semánticas del CP1 (Apéndice IX).

$\vdash$ $a$, $b$, $a$ (1.

$\vdash$ $a$, $c$, $b$, $c$, $a$, $b$, $c$ (2.

$\vdash$ $a$, $d$, $b$, $a$, $b$, $d$ (8.

$\vdash$ $b$, $a$, $a$, $b$ (28.

$\vdash$ $a$, $a$ (31.

$\vdash$ $a$, $a$ (41.

$\vdash$ $f(d)$, $f(c)$, $c \equiv d$ (52.

$\vdash$ $c \equiv c$ (54.

$\vdash$ $f(c)$, $\mathfrak{a}$ $f(\mathfrak{a})$ (58.

# XI

## DEFINICIONES RECURSIVAS (DEDEKIND 1888)

En el Capítulo 2.3 expliqué lo que Dedekind llama un *sistema simplemente infinito ordenado por una aplicación*. En el presente apéndice, una expresión de la forma $\langle\mathscr{A},\phi,\mathrm{a}\rangle$ designa un sistema simplemente infinito $\mathscr{A}$, ordenado por la aplicación $\phi$ y con elemento básico a. Si $m$, $n \in \mathscr{A}$ y $m > n$ ($m$ es mayor que $n$) en el sentido definido en la p. 158, digo también que $n < m$ ($n$ es menor que $m$). Digo, asímismo, que $m$ es *el primer* elemento de $\mathscr{A}$ que satisface una condición dada si $m \in \mathscr{A}$ y satisface dicha condición, pero ésta no es satisfecha por ningún $n \in \mathscr{A}$ que sea menor que $m$. Si $\chi$ y $\psi$ son dos aplicaciones, escribo $\psi\chi$ para referirme a la aplicación compuesta $\psi \circ \chi$.

Dedekind fue quizás el primero en entender que sólo es legítimo dar definiciones recursivas o "por inducción" si se ha demostrado que tales definiciones caracterizan en forma inequívoca aquello que se pretende definir con ellas. Su Teorema de la Definición por Inducción (TDI) responde justamente a esta necesidad. En el Apéndice VIII lo enuncié de un modo apropiado al contexto. El enunciado siguiente se ajusta mejor al texto de Dedekind:

TDI Sea $\Omega$ un sistema cualquiera. Sea $\langle\mathscr{N},\sigma,\mathbf{1}\rangle$ un sistema simplemente infinito. Si $\theta$ es una aplicación de $\Omega$ en $\Omega$ y $\omega$ es un elemento determinado de $\Omega$, hay una y sólo una aplicación $\psi: \mathscr{N} \rightarrow \Omega$ que satisface las condiciones siguientes:

I. $\psi(\mathbf{1}) = \omega$

II. Para todo $n \in \mathscr{N}$, $\psi\sigma(n) = \theta\psi(n)$.

La demostración del TDI se basa en el siguiente lema:

$\Lambda$ Sea $\Omega$ un sistema cualquiera y $n$ un elemento del sistema simplemente infinito $\langle\mathscr{N},\sigma,\mathbf{1}\rangle$. Sea $Z_n$ al sistema formado por todos los elementos de $\mathscr{N}$ iguales o menores que $n$. Si $\theta$ es una aplicación de $\Omega$ en $\Omega$ y $\omega$ es un elemento determinado de $\Omega$, hay una y sólo una aplicación $\psi_n: Z_n \rightarrow \Omega$ que satisface las condiciones siguientes:

($\alpha$) $\psi_n(\mathbf{1}) = \omega$

($\beta$) Para todo $t < n$, $\psi_n\sigma(t) = \theta\psi_n(t)$.

El lema $\Lambda$ se demuestra por inducción. Evidentemente, hay una y sólo una aplicación $\psi_{\mathbf{1}}$ de $Z_{\mathbf{1}} = \{\mathbf{1}\}$ en $\Omega$ tal que $\psi_{\mathbf{1}}(\mathbf{1}) = \omega$; $\psi_{\mathbf{1}}$ satisface trivialmente la condición ($\beta$) para $n = \mathbf{1}$, pues no hay un número $t < \mathbf{1}$. Por lo tanto, $\Lambda$ vale para $n = \mathbf{1}$. Supongamos ahora que $\Lambda$ vale para $n = q$. Demostraré que en tal caso también vale para $n = \sigma(q)$. Por hipótesis, existe una única aplicación $\psi_q$: $Z_q \to \Omega$ que cumple las condiciones ($\alpha$) y ($\beta$). Hay, por lo tanto, en $\Omega$ un único objeto $\upsilon = \theta\psi_q(q)$. Por consiguiente, la aplicación $\psi_{\sigma(q)}$: $Z_{\sigma(q)} \to \Omega$, definida por $x \mapsto \psi_q(x)$ (si $x \in Z_q$) y $\sigma(q) \mapsto \upsilon$, existe y satisface las condiciones ($\alpha$) y ($\beta$), y es claramente la única aplicación de $Z_{\sigma(q)}$ en $\Omega$ que las satisface. Por lo tanto, $\Lambda$ vale para cualquier número $n$.

Una vez asegurado el lema $\Lambda$, la demostración del TDI es fácil. Tenemos que para cada $n \in \mathcal{N}$ hay una única $\psi_n$: $Z_n \to \Omega$ que cumple las condiciones del lema. Por lo tanto, la aplicación $\psi$: $\mathcal{N} \to \Omega$, definida por $\psi(n) = \psi_n(n)$ para cada $n \in \mathcal{N}$, existe y cumple las condiciones I y II del TDI.[30] Supongamos ahora que $\chi$: $\mathcal{N} \to \Omega$ también cumple ambas condiciones pero $\chi \neq \psi$. Sea $\mu$ el primer elemento de $\mathcal{N}$ tal que $\psi(\mu) \neq \chi(\mu)$. Como $\chi$ y $\psi$ cumplen la condición I, es claro que $\mu > \mathbf{1}$. Hay, pues, un número $\nu$ tal que $\mu = \sigma(\nu)$ y $\psi(\nu) = \chi(\nu)$. Como $\chi$ y $\psi$ cumplen la condición II, $\psi(\mu) = \psi\sigma(\nu) = \theta\psi(\nu) = \theta\chi(\nu) = \chi\sigma(\nu) = = \chi(\mu)$, contradiciendo la hipótesis de que $\psi(\mu) \neq \chi(\mu)$. Por lo tanto, no hay una $\chi$: $\mathbb{N} \to \Omega$, diferente de $\psi$, que cumpla las condiciones del TDI.[31]

Mediante el TDI Dedekind puede demostrar que todos los sistemas simplemente infinitos son isomórficos. Sean $\langle\mathcal{N},\sigma,\mathbf{1}\rangle$ y $\langle\mathcal{M},\tau,\mathbf{0}\rangle$ dos sistemas simplemente infinitos. Se trata de probar que hay un isomorfismo entre ellos, esto es, una biyección $\psi$: $\mathcal{N} \to \mathcal{M}$, con inversa $\psi^{-1}$, tal que $\psi(\mathbf{1}) = \mathbf{0}$, $\tau\psi = \psi\sigma$ y $\tau\psi^{-1} = \psi^{-1}\sigma$. Según el TDI hay una y sólo una aplicación $\psi$: $\mathcal{N} \to \mathcal{M}$ tal que $\psi(\mathbf{1}) = \mathbf{0}$ y $\tau\psi = \psi\sigma$. Probaremos que $\psi$ es biyectiva. Supongamos, por el contrario, que $\psi$ no es ni siquiera inyectiva. Si hay un $n \in \mathcal{N}$ tal que

30 Obsérvese que Dedekind usa aquí ya en 1888 el método de diagonalización que Cantor hará famoso con su segunda —y más conocida— prueba de la indenumerabilidad del continuo de 1890/91. Por lo demás, ese método había sido empleado antes por du Bois- Reymond (1875); *vide supra* p. 43, nota 11.

31 Como $\psi_{\sigma(n)}$ concuerda con $\psi_n$ en $Z_n$, para cada $n \in \mathbb{N}$, resulta que $\psi_n$ es precisamente la restricción de $\psi$ a $Z_n$.

$n \neq \mathbf{1}$ y $\psi(n) = \mathbf{0}$, hay un $m \in \mathcal{N}$ tal que $n = \sigma(m)$ y $\mathbf{0} = \psi\sigma(m) = \tau\psi(m) \in \tau(\mathcal{M})$, lo cual es imposible. Sea pues $m > \mathbf{1}$ el primer elemento de $\mathcal{N}$ tal que $\psi(m) = \psi(n)$ para un $n \in \mathcal{N}$ tal que $n \neq m$. Sea $h = \sigma^{-1}(m)$ y $k = \sigma^{-1}(n)$. Como $k \neq h < m$, $\psi(h) \neq \psi(k)$ por hipótesis. Como $\tau$ es inyectiva, $\tau\psi(h) \neq \tau\psi(k)$. Pero $\tau\psi(h) = \psi\sigma(h) = \psi\sigma(k) = \tau\psi(k)$. ¡Contradicción! Por lo tanto, $\psi$ es inyectiva. Supongamos ahora que $\psi$, aunque inyectiva, no es biyectiva. Entonces hay elementos de $\mathcal{M}$ que no pertenecen a $\psi(\mathcal{N})$. Sea $a$ el primer elemento de $\mathcal{M}$ tal que $a \notin \psi(\mathcal{N})$. Como $\mathbf{0} = \psi(\mathbf{1})$, $a \neq \mathbf{0}$. Por lo tanto, hay un $b \in \mathcal{M}$ tal que $a = \tau(b)$ y hay un $c \in \mathcal{N}$ tal que $b = \psi(c)$. Pero entonces $a = \tau\psi(c) = \psi\sigma(c)$, de modo que $a \in \psi(\mathcal{N})$. ¡Contradicción! Por lo tanto, $\psi$ es biyectiva. Sea, pues, $\psi^{-1}$: $\mathcal{M} \to \mathcal{N}$ la biyección inversa. Tenemos que $\psi^{-1}(\mathbf{0}) = \mathbf{1}$. Además, $\psi^{-1}\tau = \psi^{-1}\tau(\psi\psi^{-1}) = \psi^{-1}(\tau\psi)\psi^{-1} = \psi^{-1}(\psi\sigma)\psi^{-1} = (\psi^{-1}\psi)\sigma\psi^{-1} = \sigma\psi^{-1}$. Esto completa la demostración de que $\psi$ es un isomorfismo. Se observará que $\psi^{-1}$ es precisamente la aplicación de $\mathcal{M}$ en $\mathcal{N}$ cuya existencia y unicidad establece el TDI.

La unicidad de $\psi$ y $\psi^{-1}$ significa que cada elemento de $\mathcal{M}$ puede tomarse como representante canónico de un determinado elemento de $\mathcal{N}$ y viceversa. Por esta razón, se justifica plenamente sostener con Dedekind que cada sistema simplemente infinito puede servirnos de "serie numérica" lo mismo que cualquier otro (cf. Capítulo 2.3). En lo que resta de este apéndice, hablaré sólo de una tal estructura fija $\langle\mathbb{N},\sigma,\mathbf{0}\rangle$, que elijo como representativa.[32] Siguiendo a Dedekind, definiré la adición, la multiplicación y la exponenciación en esta estructura y demostraré que la adición y la multiplicación así definidas tienen las propiedades tradicionales: conmutatividad, asociatividad, distributividad de la multiplicación con respecto a la adición. Las demostraciones se hacen por inducción. Para facilitar la lectura sigo en todas el mismo esquema: enuncio una tesis [T] concerniente a un número indeterminado $n$; establezco la base de la inducción [B]: la tesis es verdadera si $n = \mathbf{0}$; y doy el llamado "paso inductivo" [P], esto es, pruebo que si la tesis vale para un dado $n \in \mathbb{N}$ también vale para el elemento siguiente $\sigma(n)$.

[32] Conforme a la práctica matemática actual, designo con $\mathbf{0}$ el elemento básico de la "serie numérica" representativa (contamos: "cero, uno, dos, . . . "). Como se verá, ello permite definir elegantemente la adición de modo que la operación de sumar el elemento básico ($x \mapsto x + \mathbf{0}$) coincida con la identidad $\mathcal{I}_{\mathbb{N}}: \mathbb{N} \to \mathbb{N}; x = x{+}1$.

*Adición*. En virtud del TDI existe, para cada $n \in \mathbb{N}$, una y sólo una aplicación $\sigma_n$: $\mathbb{N} \to \mathbb{N}$, tal que $\sigma_n(\mathbf{0}) = n$ y $\sigma_n\sigma = \sigma\sigma_n$. En particular, $\sigma_{\mathbf{0}}$ es la identidad en $\mathbb{N}$, $\mathscr{I}_{\mathbb{N}}$: $x \mapsto x$, pues ésta es la única aplicación de $\mathbb{N} \to \mathbb{N}$ que conmuta con σ y asigna el valor **0** al argumento **0**. La adición queda entonces definida inequívocamente por la condición: $n + m = \sigma_m(n)$. Probaremos por inducción que, para todo $m, n \in \mathbb{N}$, $\sigma(n + m) = n + \sigma(m)$.

[T] $\sigma\sigma_m(n) = \sigma_{\sigma(m)}(n)$.

[B] Es claro que $\sigma_{\sigma(m)}(\mathbf{0}) = \sigma(m) = \sigma\sigma_m(\mathbf{0})$.

[P] Si $\sigma(n + m) = \sigma\sigma_m(n) = \sigma_{\sigma(m)}(n) = n + \sigma(m)$, entonces $\sigma(\sigma(n) + m) = \sigma\sigma_m\sigma(n) = \sigma\sigma\sigma_m(n) = \sigma\sigma_{\sigma(m)}(n) = \sigma_{\sigma(m)}\sigma(n) = (\sigma(n) + \sigma(m))$.

Por lo tanto, $\sigma(n + m) = n + \sigma(m)$, para todo $n$, cualquiera que sea $m$. Si ponemos $m = \mathbf{0}$, comprobamos que $\sigma_{\sigma(\mathbf{0})} = \sigma\sigma_{\mathbf{0}} = \sigma$.

*La adición es asociativa*.

[T] Si $a$, $b$ y $n$ son cualesquiera elementos de $\mathbb{N}$, entonces $(a + b) + n = a + (b + n)$.

[B] $(a + b) + \mathbf{0} = a + b = a + (b + \mathbf{0})$

[P] Si $(a + b) + n = a + (b + n)$, entonces $(a + b) + \sigma(n) = \sigma((a + b) + n) = \sigma(a + (b + n)) = a + \sigma(b + n) = a + (b + \sigma(n))$.

Por lo tanto, la adición es asociativa.

*La adición es conmutativa*.

[T] Si $a$ y $n$ son cualesquiera elementos de $\mathbb{N}$, $a + n = n + a$.

[B] $a + \mathbf{0} = \sigma_{\mathbf{0}}(a) = a = \sigma_a(\mathbf{0}) = \mathbf{0} + a$.

[P] Si $a + n = n + a$, esto es, si $\sigma_n\sigma_a(\mathbf{0}) = \sigma_a\sigma_n(\mathbf{0})$, tenemos que $a + \sigma(n) = \sigma\sigma_n\sigma_a(\mathbf{0}) = \sigma\sigma_a\sigma_n(\mathbf{0}) = \sigma_a\sigma\sigma_n(\mathbf{0}) = \sigma(n) + a$.

Por lo tanto, la adición es conmutativa.[33]

[33] Como Dedekind cuenta "uno, dos, tres, . . . ", tiene que definir la adición de modo que el elemento básico sumado a cualquier $n$ no dé $n$ sino el siguiente de $n$. Sea pues el sistema simplemente infinito $\mathbb{N}$ ordenado por la aplicación σ la serie numérica representativa elegida por Dedekind, y designemos con **1** su elemento básico. Dedekind considera para cada $n \in \mathbb{N}$ la única aplicación $\varphi_n$: $\mathbb{N} \to \mathbb{N}$, tal que $\varphi_n(\mathbf{1}) = \sigma(n)$ y $\varphi_n\sigma = \sigma\varphi_n$. Obviamente, $\varphi_{\mathbf{1}} = \sigma$. La adición se define por la condición $m + n = \varphi_n(m)$, la cual implica que (i) $m + \mathbf{1} = \sigma(m)$ y (ii) $\sigma(m + n) = m + \sigma(n)$. La conmutatividad y asociatividad de la adición pueden establecerse entonces como arriba.

*Multiplicación*. Para definir la multiplicación, conviene tener presente que, en virtud del TDI, si $n \in \mathbb{N}$ hay una y una sola aplicación $\chi_n: \mathbb{N} \rightarrow \mathbb{N}$ tal que $\chi_n(\mathbf{0}) = \mathbf{0}$ y $\sigma_n\chi_n = \chi_n\sigma$. Obsérvese, en particular que, para todo $n \in \mathbb{N}$, $\chi_{\mathbf{0}}(n) = \mathbf{0}$ y $\chi_{\sigma(\mathbf{0})}(n) = n$.[34] La multiplicación queda inequívocamente definida por la condición $m \times n = \chi_n(m)$. En virtud de la observación recién hecha, $n \times \mathbf{0} = \mathbf{0} = \mathbf{0} \times n$ y $n \times \sigma(\mathbf{0}) = n$.

*La multiplicación es distributiva.*

[T] $n \times (a + b) = (n \times a) + (n \times b)$.

[B] $\mathbf{0} \times (a + b) = \mathbf{0} = (\mathbf{0} \times a) + (\mathbf{0} \times b)$.

[P] Si $n \times (a + b) = (n \times a) + (n \times b)$, $\sigma(n) \times (a + b) = \chi_{a+b}\sigma(n) = \sigma_{a+b}\chi_{a+b}(n) = (n \times (a + b)) + a + b = (n \times a) + (n \times b) + a + b = (n \times a) + a + (n \times b) + b =, \sigma_a\chi_a(n) + \sigma_b\chi_b(n) = \chi_a\sigma(n) + \chi_b\sigma(n) = (\sigma(n) \times a) + (\sigma(n) \times b)$.

Por lo tanto, la multiplicación es distributiva (con respecto a la adición).

*La multiplicación es conmutativa.*

[T] $n \times a = a \times n$.

[B] $\mathbf{0} \times a = \mathbf{0} = a \times \mathbf{0}$.

[P] Si $n \times a = a \times n$, $\sigma(n) \times a = \chi_a\sigma(n) = \sigma_a\chi_a(n) = (n \times a) + a = (a \times n) + (a \times \sigma(\mathbf{0})) = a \times (n + \sigma(\mathbf{0})) = a \times \sigma(n)$.

Por lo tanto, la multiplicación es conmutativa.

*La multiplicación es asociativa.*

[T] $n \times (a \times b) = (n \times a) \times b$.

[B] $\mathbf{0} \times (a \times b) = \mathbf{0} = (\mathbf{0} \times a) \times b$

[P] Si $n \times (a \times b) = (n \times a) \times b$, $\sigma n \times (a \times b) = \chi_{a \times b}\sigma(n) = \sigma_{a \times b}\chi_{a \times b}(n) = (n \times (a \times b)) + (a \times b) = ((n \times a) \times b) + (a \times b) = (b \times (n \times a)) + (b \times a) = b \times ((n \times a) + a) = b \times ((a \times n) + (a \times \sigma(\mathbf{0})) = b \times (a \times (n + \sigma(\mathbf{0}))) = b \times (a \times \sigma(n)) = ((\sigma(n) \times a) \times b$.

Por lo tanto, la multiplicación es asociativa.

34 Por una parte, $\chi_{\mathbf{0}}\sigma(n) = \sigma_{\mathbf{0}}\chi_{\mathbf{0}}(n) = \chi_{\mathbf{0}}(n) + \mathbf{0} = \chi_{\mathbf{0}}(n) = \mathbf{0}$, puesto que $\chi_{\mathbf{0}}(\mathbf{0}) = \mathbf{0}$. Por otra parte, $\chi_{\sigma(\mathbf{0})}(\sigma(\mathbf{0})) = \sigma(\mathbf{0})$ y si $\chi_{\sigma(\mathbf{0})}(n) = n$, entonces $\chi_{\sigma(\mathbf{0})}\sigma(n) = \sigma_{\sigma(\mathbf{0})}\chi_{\sigma(\mathbf{0})}(n) = \sigma\sigma_{\mathbf{0}}\chi_{\sigma(\mathbf{0})}(n) = \sigma\chi_{\sigma(\mathbf{0})}(n) = \sigma(n)$.

*Exponenciación*. Para definir la exponenciación, consideramos que, en virtud del TDI, si $n \in \mathbb{N}$ hay una y sólo una aplicación $\varepsilon_n : \mathbb{N} \to \mathbb{N}$ tal que $\varepsilon_n(\mathbf{0}) = \sigma(\mathbf{0})$ y $\chi_n\varepsilon_n = \varepsilon_n\sigma$. Definimos: $m^n = \varepsilon_n(m)$. El lector puede ejercitarse en probar que $m^0 = \sigma(\mathbf{0})$, $m^p m^q = m^{p+q}$ y $(m^p)^q = m^{pq}$, cualesquiera que sean $m, p, q \in \mathbb{N}$.

## XII

## EXTENSIÓN Y RECORRIDO (FREGE 1891, 1893)

Según Frege (1884), cada número natural es la extensión de un concepto, a saber, el concepto mediante el cual se piensa la propiedad de ser *equinumeroso con cierto concepto* (éste puede ser cualquiera de los conceptos a que —en sentido fregeano— corresponde el número en cuestión). Para Frege, las extensiones de conceptos son objetos, que pueden naturalmente caer bajo otros conceptos —por ejemplo, la extensión del concepto *equinumeroso con el concepto 'día de la semana'* (esto es, el número 7) y la extensión del concepto *equinumeroso con el concepto 'mes del año'* (esto es, el número 12) caen ambas bajo el concepto *número finito*. Como Frege (1884, p. 80n.) dice que da por sabido lo qué es la extensión (*Umfang*) de un concepto, en mi presentación de su filosofía de la aritmética en la Sección 2.3.3 asigné a ese término su significado tradicional, con arreglo al cual la extensión de un concepto es la suma o agregado de todas las cosas que caen bajo ese concepto.

En su versión original, la escritura conceptual BS no incluía un medio específico para nombrar la extensión de un concepto. Frege (1891) remedia esta deficiencia introduciendo lo que luego describe como "uno de los enriquecimientos más fecundos (*folgenreichsten*) de mi escritura conceptual" (1893, p. 14). De ahí en adelante, la notación $\grave{\varepsilon}F(\varepsilon)$ designa la extensión —

en el sentido de Frege— del concepto expresado por la letra *F*.[35] Con todo, las explicaciones que rodean la introducción del nuevo recurso expresivo parecen indicar que, en su obra madura, Frege no entiende el término ‘extensión de un concepto’ en la acepción tradicional arriba señalada.

Como indiqué en el Capítulo 2.2, desde 1891 Frege entiende que los *conceptos* constituyen una especie del género *función*. Un concepto es una función que asigna a cada objeto del universo uno de los dos valores v (“lo verdadero”) o f (“lo falso”). Frege supone que cada función esta asociada a un objeto característico, que llamaré su *recorrido* (Frege dice *Wertverlauf*, literalmente, “recorrido del valor”). Frege no explica en qué consiste el recorrido de una función. Se limita a decir que “emplea universalmente las palabras ‘la función $\Phi(\xi)$ tiene el mismo *recorrido* que la función $\Psi(\xi)$’ como sinónimas de las palabras ‘las funciones $\Phi(\xi)$ y $\Psi(\xi)$ tienen siempre valores iguales para argumentos iguales’” (1893, p. 7). La notación ἐ$\Phi(\varepsilon)$ se introduce precisamente para designar el recorrido de la función $\Phi$ (1891, p. 10).[36]

Frege adopta la siguiente convención léxica: “Podemos designar como extensión de un concepto (*Begriffsumfang*) al recorrido de una función cuyo valor para cada argumento es un valor veritativo (*Wahrheitswert*)”, esto es, uno de los objetos v o f (1891, p. 16). En virtud de ella, si la función $\Phi$ es un concepto, la expresión ‘ἐ$\Phi(\varepsilon)$’ denota su extensión —en el nuevo sentido fregeano— mas no parece que ésta pueda identificarse con la colección de todas las cosas que caen bajo ese concepto.

Las funciones fregeanas son, como sabemos, aplicaciones del universo de objetos en sí mismo. La equivalencia ἐ$\Phi(\varepsilon)$ = ἐ$\Psi(\varepsilon) \leftrightarrow \forall x(\Phi(x) \leftrightarrow \Psi(x))$ sugiere, pues, que el recorrido de la función $\Phi$ es lo que hoy llamaríamos su *grafo*, esto es, el conjunto de todos los pares $\langle x, \Phi(x)\rangle$ que se forman toman-

35 Hay una ironía involuntaria en el epíteto *folgenreichste* —literalmente, ‘más rico en consecuencias’— aplicado a la nueva notación pues gracias al axioma introducido por Frege para regular su uso resulta que *todas* las aseveraciones expresables en la escritura BS son derivables en el sistema fregeano.

36 La notación se extiende fácilmente a lo que Frege llama ‘funciones con dos (o más) argumentos’, tal como la suma, designada por la expresión insaturada $\xi + \zeta$ (en que las letras griegas marcan los “huecos”). Si fijamos uno de los dos argumentos, poniendo, digamos, $\zeta = 3$, obtenemos una función de un argumento, cuyo recorrido es ἐ$(\varepsilon + 3)$. Entonces, la expresión insaturada ‘ἐ$(\varepsilon + \zeta)$’ designa una función de un argumento, cuyo recorrido es ἀἐ$(\varepsilon + \alpha)$.

do cada argumento $x$ de la función $\Phi$ junto con el valor correspondiente $\Phi(x)$. Los argumentos de una función fregeana son todos los objetos del universo. Si la función es un concepto, los valores admisibles son sólo los dos valores veritativos. Por lo tanto, el grafo de un concepto consta de dos clases de pares, aquellos cuyo segundo miembro es v y aquellos cuyo segundo miembro es f, y divide exhaustiva y exclusivamente los objetos del universo en dos grupos, aquellos que forman el primer miembro en un par de la primera clase y aquellos que forman el primer miembro de un par de la segunda clase. Si entendemos, pues, que el recorrido de una función fregeana es precisamente su grafo, entonces la extensión fregeana $\dot{\varepsilon}F(\varepsilon)$ del concepto expresado por la letra $F$ puede verse como una partición del universo —lo que propongo llamar una *cortadura de Frege* (*Frege-Schnitt* en alemán, *Frege-cut* en inglés)— en la que cada objeto $x$ figura marcado con uno de los dos valores veritativos. Por otra parte, dos conceptos $F$ y $G$ tienen el mismo recorrido si y sólo si los objetos que caen bajo $F$ son precisamente los mismos que caen bajo $G$. (Como es obvio, si $\Phi$ y $\Psi$ son conceptos, $\dot{\varepsilon}\Phi(\varepsilon) = \dot{\varepsilon}\Psi(\varepsilon) \leftrightarrow \forall x(\Phi(x) \leftrightarrow \Psi(x)) \leftrightarrow \forall x(\Phi(x) = \text{v} \leftrightarrow \Psi(x) = \text{v})$.)

Frege (1893) utiliza eficazmente la notación $\dot{\varepsilon}F(\varepsilon)$ para expresar en BS la definiciones de conceptos aritméticos fundamentales que había formulado en alemán en 1884. Las nuevas definiciones son sumamente ingeniosas y es una lástima que no pueda comentarlas aquí. La nueva notación figura sólo en dos de las "leyes lógicas fundamentales" en que descansa el sistema deductivo de Frege. Aquí sólo nos interesa la primera:[37]

(V) $$(\dot{\varepsilon}f(\varepsilon) = \dot{\alpha}g(\alpha)) = (\neg\!\overset{\mathfrak{u}}{\smile}\!\!-\, f(\mathfrak{u}) = g(\mathfrak{u}))$$

[37] La otra ley fundamental en que figura la notación $\dot{\varepsilon}\Phi(\varepsilon)$ es esta:

(VI) $$\vdash a = \backslash\dot{\varepsilon}(a = \varepsilon)$$

Ella incorpora formalmente al sistema la notación '$\backslash\xi$' que Frege 1893, p. 19, explica informalmente en estos términos:

> Distinguimos dos casos:
>
> 1) Si, para el argumento [de la función $\backslash\xi$] existe un objeto $\Delta$ tal que $\backslash\dot{\varepsilon}(\Delta = \varepsilon)$ sea el argumento, entonces $\Delta$ mismo es el valor de la función $\backslash\xi$.
> 2) Si para el argumento [de la función $\backslash\xi$] no existe ningún objeto $\Delta$ tal que $\backslash\dot{\varepsilon}(\Delta = \varepsilon)$ sea el argumento, entonces el argumento mismo es el valor de la función $\backslash\xi$.

Ella repite formalmente y a título de axioma la indicación arriba citada acerca de las condiciones en que los recorridos de dos conceptos son iguales. Este Axioma V de Frege 1893 lleva derechamente a la paradoja de Russell. La siguiente derivación fue publicada por Frege (1903, Nachwort) poco después de que Russell le escribió al respecto. Para facilitar la lectura, "traduzco" las fórmulas de Frege al simbolismo lógico actual, enriquecido con la notación fregeana para nombrar recorridos. En particular, reemplazo el signo fregeano de identidad '=' por nuestro signo '=' cuando figura entre nombres de objetos, y por nuestro signo de equivalencia '$\leftrightarrow$' cuando figura entre oraciones. Apelaré asímismo a reglas de inferencia y teoremas familiares de la lógica proposicional y predicativa, todos los cuales pueden por cierto validarse en el sistema de Frege (sin recurrir al Axioma V).

"Traducido", el Axioma V reza así:

$$\forall G \forall F(\text{ἐ}F\varepsilon = \text{ἀ}G\alpha \leftrightarrow \forall u(Fu \leftrightarrow Gu)) \tag{V}$$

Con nuestros símbolos de conjunción y cuantificación existencial, combinados con la notación $\text{ἐ}\Phi(\varepsilon)$, se puede expresar perspicuamente, mediante la fórmula $\exists G(\text{ἐ}G\varepsilon = \Delta \wedge \neg G\Delta)$, que $\Delta$ es una clase que no pertenece a sí misma. Frege (1903, p. 256) emplea la fórmula $\top\!\overset{\mathfrak{G}}{\smile}\!\top\!\!\begin{array}{l}\mathfrak{G}(\Delta)\\ \text{ἐ}(-\mathfrak{G}(\varepsilon)) = \Delta\end{array}$ la cual equivale lógicamente a la anterior, aunque sería más exacto traducirla así: $\neg\forall G(\text{ἐ}G\varepsilon = \Delta \rightarrow G\Delta)$. La clase de todas las clases que no se pertenecen a sí mismas debe, pues, designarse con el nombre del recorrido de la función $\neg\forall G(\text{ἐ}G\varepsilon = \xi \rightarrow G\xi)$, vale decir, $\text{ἐ}(\neg\forall G(\text{ἐ}G\varepsilon = \varepsilon \rightarrow G\varepsilon))$. En aras de la brevedad, denotaré este objeto con la constante $k$.[38] Así, por definición:

---

Según esto, $\backslash\text{ἐ}(\Delta = \varepsilon) = \Delta$ es lo verdadero y '$\backslash\text{ἐ}\Phi(\varepsilon)$' denota el objeto que cae bajo el concepto $\Phi(\xi)$ si $\Phi(\xi)$ es un concepto bajo el cual cae un y sólo un objeto. En todos los demás casos, '$\backslash\text{ἐ}\Phi(\varepsilon)$' denota lo mismo que '$\text{ἐ}\Phi(\varepsilon)$'.

[38] Frege también usa una abreviatura para nombrar la clase de todas las clases que pertenecen a sí mismas, pero, como cabía esperar, es mucho más rebuscada que la adoptada aquí. Usa la letra A invertida, creando así —con otro significado— el símbolo moderno del cuantificador universal ($\forall$).

$$\dot{\varepsilon}(\neg\forall G(\dot{\varepsilon}G\varepsilon = \varepsilon \rightarrow G\varepsilon)) = k \tag{1}$$

Si en el Axioma V ejemplificamos las variables predicativas $F$ y $G$, respectivamente con las funciones $H\xi$ y $\neg\forall G(\dot{\varepsilon}G\varepsilon = \xi \rightarrow G\xi)$, deducimos la proposición siguiente:

$$\dot{\varepsilon}H\varepsilon = \dot{\varepsilon}(\neg\forall G(\dot{\varepsilon}G\varepsilon = \varepsilon \rightarrow G\varepsilon)) \leftrightarrow \forall u(Hu \leftrightarrow \neg\forall G(\dot{\varepsilon}G\varepsilon = u \rightarrow Gu)) \tag{2}$$

Para facilitar la comprensión de los pasos siguientes abrevio con $\forall u\Psi(u)$ la subfórmula a la derecha del primer signo $\leftrightarrow$. La convención (1) me permite reemplazar con $\dot{\varepsilon}H\varepsilon = k$ la subfórmula a la izquierda de dicho signo. Usando estas abreviaturas, (2) se escribe:

$$\dot{\varepsilon}H\varepsilon = k \leftrightarrow \forall u\Psi(u) \tag{2a}$$

(2a) implica que $\dot{\varepsilon}H\varepsilon = k \rightarrow \forall u\Psi(u)$. De esta fórmula —utilizando el esquema (12) del Apéndice XIII (p. 518)— se deduce la fórmula siguiente:

$$\forall u(\dot{\varepsilon}H\varepsilon = k \rightarrow \Psi(u)) \tag{3}$$

Ejemplificando la variable individual $u$ con el objeto $k$, obtenemos:

$$\dot{\varepsilon}H\varepsilon = k \leftrightarrow \Psi(k) \tag{4a}$$

Si recordamos que $\Psi(u)$ representa el alcance del cuantificador $\forall u$ en (2), vemos que (4a) abrevia la fórmula siguiente:

$$\dot{\varepsilon}H\varepsilon = k \leftrightarrow (Hk \leftrightarrow \neg\forall G(\dot{\varepsilon}G\varepsilon = k \rightarrow Gk)) \tag{4}$$

La proposición (4) implica que:[39]

$$\neg\forall G(\dot{\varepsilon}G\varepsilon = k \rightarrow Gk) \rightarrow (\dot{\varepsilon}H\varepsilon = k \rightarrow Hk) \tag{5}$$

[39] (4) es una función veritativa de los componentes $\dot{\varepsilon}H\varepsilon = k$, $Hk$ y $\neg\forall G(\dot{\varepsilon}G\varepsilon = k \rightarrow Gk)$, que representaremos con $P$, $Q$ y $R$, respectivamente. Un breve cálculo permite comprobar que $(P \leftrightarrow (Q \rightarrow R)) \rightarrow (R \rightarrow (P \rightarrow Q))$ es una tautología.

Como $H$ es una variable predicativa arbitraria que no figura en las premisas iniciales de nuestra deducción, es lícito generalizar:

$$\forall H(\neg\forall G(\dot{\varepsilon}G\varepsilon = k \rightarrow Gk) \rightarrow (\dot{\varepsilon}H\varepsilon = k \rightarrow Hk)) \qquad (6)$$

Invocando los análogos de segundo orden de los esquemas (12) y (7) del Apéndice XIII, derivamos de (6) la proposición (7), que Frege deriva directamente de (5):

$$\neg\forall G(\dot{\varepsilon}G\varepsilon = k \rightarrow Gk) \rightarrow \forall G(\dot{\varepsilon}G\varepsilon = k \rightarrow Gk) \qquad (7)$$

Por otra parte, mediante una simple aplicación del Axioma IIb de Frege 1893 (que es la versión de segundo orden del Axioma 58 de Frege 1879, reproducido arriba en la p. 503), obtenemos:

$$\forall G(\dot{\varepsilon}G\varepsilon = k \rightarrow Gk) \rightarrow (\dot{\varepsilon}H\varepsilon = k \rightarrow Hk) \qquad (8)$$

Reemplazando la función arbitraria $H\xi$ con $\neg\forall G(\dot{\varepsilon}G\varepsilon = \xi \rightarrow G\xi)$, deducimos:

$$\forall G(\dot{\varepsilon}G\varepsilon = k \rightarrow Gk) \rightarrow$$
$$(\dot{\varepsilon}(\neg\forall G(\dot{\varepsilon}G\varepsilon = \varepsilon \rightarrow G\varepsilon)) = k \rightarrow \neg\forall G(\dot{\varepsilon}G\varepsilon = k \rightarrow Gk)) \qquad (9)$$

Utilizando la convención (1), (9) se puede abreviar así:

$$\forall G(\dot{\varepsilon}G\varepsilon = k \rightarrow Gk) \rightarrow (k = k \rightarrow \neg\forall G(\dot{\varepsilon}G\varepsilon = k \rightarrow Gk)) \qquad (10)$$

Por lo tanto,

$$\forall G(\dot{\varepsilon}G\varepsilon = k \rightarrow Gk) \rightarrow \neg\forall G(\dot{\varepsilon}G\varepsilon = k \rightarrow Gk)) \qquad (11)$$

Sustituyendo $P$ por $\forall G(\dot{\varepsilon}G\varepsilon = k \rightarrow Gk)$ en la tautología $(P \rightarrow \neg P) \rightarrow \neg P$ y aplicando la regla *modus ponens* deducimos:

$$\neg\forall G(\dot{\varepsilon}G(\varepsilon) = k \rightarrow Gk) \qquad (12)$$

De (12) y (7), por *modus ponens*, se deduce:

$$\forall G(\dot{\varepsilon}G\varepsilon = k \rightarrow Gk) \tag{13}$$

que combinada con (12) da

$$\forall G(\dot{\varepsilon}G\varepsilon = k \rightarrow Gk) \wedge \neg\forall G(\dot{\varepsilon}G\varepsilon = k \rightarrow Gk) \tag{14}$$

La contradicción (14) es una consecuencia lógica del Axioma V de Frege combinado con sus demás axiomas y reglas de inferencia. (14) implica, a su vez, todas las proposiciones enunciables en BS.

En el mismo Epílogo que contiene la deducción anterior, Frege sugiere un remedio. Descompone el Axioma V en dos implicaciones:

$$\forall G\forall F(\forall u(Fu \leftrightarrow Gu) \rightarrow \dot{\varepsilon}F\varepsilon = \dot{\alpha}G\alpha) \tag{Va}$$

$$\forall G\forall F(\dot{\varepsilon}F\varepsilon = \dot{\alpha}G\alpha \rightarrow \forall u(Fu \leftrightarrow Gu)) \tag{Vb}$$

Según él, la paradoja de Russell es una consecuencia de (Vb). El remedio propuesto consiste en reemplazarla por

$$\forall G\forall F(\dot{\varepsilon}F\varepsilon = \dot{\alpha}G\alpha \rightarrow \forall u(u \neq \dot{\varepsilon}F\varepsilon \rightarrow (Fu \leftrightarrow Gu))) \tag{Vb$'$}$$

Años más tarde, Leśniewski (en 1938; *vide* Sobociński 1949), Quine (1955) y Geach (1956) demostraron que el sistema así modificado implica una contradicción si suponemos que existen por lo menos dos objetos.[40] Pero Frege

[40] Resnik 1980, pp. 214ss., explica bien este asunto. Al comienzo de su exposición Resnik observa que (Vb) es manifiestamente incompatible con el Teorema de Cantor (cuya demostración —como vimos en la p. 50— sugirió a Russell su paradoja). En efecto, podemos entender que el Axioma V postula la existencia de una aplicación del reino de los conceptos en el universo de los objetos, y, en tal caso, (Vb) dice que dicha aplicación es inyectiva. Según Resnik, ello contradice el Teorema de Cantor, conforme al cual "hay más conjuntos de objetos que objetos" (1980, p. 214). Esta observación me parece muy confusa. Los conjuntos de objetos, en el sentido de Cantor, también son objetos, no conceptos, y lo que el Teorema de Cantor dice es que si $S$ es un conjunto *cualquiera* de objetos —sean ellos objetos individuales o conjuntos— no puede haber una aplicación inyectiva de $\mathcal{P}S$ en $S$, donde $\mathcal{P}S$ es el conjunto de *objetos* cons-

debe haber sabido que su remedio no servía, puesto que después de 1903 dejó de interesarse en la derivación de las verdades aritméticas de leyes lógicas.

## XIII

## FÓRMULAS PRENEXAS

Este es un suplemento del Apéndice IX. Me referiré específicamente a una versión del CP1= que contenga un número ilimitado de predicados *n*-ádicos, para cada entero $n \geq 0$; pero todo lo que digo puede aplicarse, *mutatis mutandis*, a otros cálculos análogos, de orden primero o superior. Recordemos que si $\alpha$ es una fórmula y $\xi$ y $\zeta$ son variables, $\alpha_{\zeta/\xi}$ es la fórmula que se obtiene al sustituir $\xi$ por $\zeta$ en todas las posiciones libres de $\xi$ en $\alpha$.

Es fácil mostrar, apelando a las reglas semánticas, que los siguientes esquemas representan fórmulas válidas:

(1) $$\alpha \leftrightarrow \neg\neg\alpha$$

(2) $$\exists\xi\alpha \leftrightarrow \neg\forall\xi\neg\alpha$$

(3) $$\forall\xi\alpha \leftrightarrow \neg\exists\xi\neg\alpha$$

(4) $$\forall\xi\neg\alpha \leftrightarrow \neg\exists\xi\alpha$$

(5) $$\exists\xi\neg\alpha \leftrightarrow \neg\forall\xi\alpha$$

Usaré en adelante la letra Q para representar indistintamente los signos de

tituido por las partes (subconjuntos) de *S*. Sólo si *presuponemos* que cada concepto determina un objeto que le corresponde en forma exclusiva —sea éste su extensión clásica, su recorrido fregeano, o lo que se quiera— podemos relacionar las correspondencias entre objetos a que se refiere el Teorema de Cantor con la correspondencia entre conceptos y objetos postulada por el Axioma V de Frege.

cuantificación $\forall$ y $\exists$. Sabemos que, si $\xi$ es una variable, $Q\xi\alpha$ es una fórmula si y sólo si $\alpha$ es una fórmula. En tal caso, $\alpha$ es el *alcance* del *cuantificador* $Q\xi$, el cual *liga* la variable $\xi$ en la posición que $\xi$ tiene dentro de él y en todas las posiciones libres que $\xi$ ocupa en $\alpha$. Entenderemos que Q representa el mismo signo de cuantificación todas las veces que ligue una misma variable en una expresión dada.

Sea $\alpha$ una fórmula en que la variable $\xi$ no ocupa ninguna posición libre. Entonces todas las fórmulas representadas por el esquema siguiente son válidas en virtud de las reglas semánticas $[I_\forall]$, $[I_\exists]$ y $[I_\leftrightarrow]$:

(6) $$Q\xi\alpha \leftrightarrow \alpha$$

Por lo tanto, $Q\xi\alpha$ y $\alpha$ se pueden sustituir mutuamente en cualquier fórmula $\beta$ en la cual una de ellas figure como subfórmula, sin que dicha sustitución afecte la verdad o la falsedad de $\beta$. Decimos, por eso, que el cuantificador $Q\xi$ está ocioso en la fórmula $Q\xi\alpha$.

Por otra parte, si $\beta$ es una fórmula cualquiera en que la variable $\zeta$ no ocupa ninguna posición libre, y ninguna posición de $\xi$ en $\beta$ cae dentro del alcance de un cuantificador que ligue a $\zeta$, es fácil comprobar que, según las reglas semánticas del CP1=,

(7) $$Q\xi\beta \leftrightarrow Q\zeta\beta_{\zeta/\xi}$$

Diré que una fórmula $\phi$ es *regular* si cumple las siguientes condiciones: (i) $\phi$ no contiene ningún cuantificador ocioso; (ii) ninguna variable ligada en $\beta$ por un cuantificador ocupa en $\beta$ posiciones donde esté libre o ligada por otro cuantificador. Obsérvese que si una fórmula es regular, también son regulares todas sus subfórmulas. Aplicando repetidamente (6) y (7) se puede construir, para cada fórmula $\psi$, una fórmula regular lógicamente equivalente a $\psi$. Hay muchos modos de hacerlo, pero es fácil prescribir un método que determine una fórmula regular única $\psi^*$ tal que $(\psi \leftrightarrow \psi^*)$ sea válida. En-

[41] Por ejemplo, se puede adoptar el método siguiente. Sea $\langle Q_1, \ldots, Q_r \rangle$ la lista de los cuantificadores no ociosos de $\psi$, en el orden de sus posiciones respectivas, y $\zeta_1, \zeta_2, \ldots$ las secuencia de las variables que no figuran en $\psi$, ordenadas según el número de palotes. Sea $\psi^*$ la fórmula resultante cuando $\psi$ es sometida a las operaciones siguientes: (i) eliminar todos los cuantificadores ociosos; (ii) reemplazar por $\zeta_i$ la variable ligada por

tonces $\psi^*$ es *la* fórmula regular *correspondiente* a $\psi$.[41]

Suponemos que en las equivalencias (8)–(14), el esquema a la izquierda del signo $\leftrightarrow$ representa una fórmula regular. Recordando que $(\alpha \rightarrow \beta) \leftrightarrow (\neg\alpha \vee \beta)$ y $(\alpha \wedge \beta) \leftrightarrow \neg(\neg\alpha \vee \neg\beta)$ son válidas, es fácil establecer la validez de:

(8) $$(Q\xi\alpha \vee \beta) \leftrightarrow Q\xi(\alpha \vee \beta)$$

(9) $$(\alpha \vee Q\xi\beta) \leftrightarrow Q\xi(\alpha \vee \beta)$$

(10) $$(Q\xi\alpha \wedge \beta) \leftrightarrow Q\xi(\alpha \wedge \beta)$$

(11) $$(\alpha \wedge Q\xi\beta) \leftrightarrow Q\xi(\alpha \wedge \beta)$$

(12) $$(\alpha \rightarrow Q\xi\beta) \leftrightarrow Q\xi(\alpha \rightarrow \beta)$$

(13) $$(\forall\xi\alpha \rightarrow \beta) \leftrightarrow \exists\xi(\alpha \rightarrow \beta)$$

(14) $$(\exists\xi\alpha \rightarrow \beta) \leftrightarrow \forall\xi(\alpha \rightarrow \beta)$$

Y si tenemos en cuenta que $(\alpha \leftrightarrow \beta)$ equivale lógicamente a $(\alpha \wedge \beta) \vee (\neg\alpha \wedge \neg\beta)$, un breve cómputo confirmará que:

(15) $$(\forall\xi\alpha \leftrightarrow \beta) \leftrightarrow \forall\xi\exists\zeta((\alpha \wedge \beta) \vee (\neg\alpha_{\zeta/\xi} \wedge \neg\beta))$$

(16) $$(\exists\xi\alpha \leftrightarrow \beta) \leftrightarrow \exists\xi\forall\zeta((\alpha \wedge \beta) \vee (\neg\alpha_{\zeta/\xi} \wedge \neg\beta))$$

donde $(Q\xi\alpha \leftrightarrow \beta)$ es regular y $\zeta$ es una variable que no figura en $\alpha$ ni en $\beta$. Como $(\beta \leftrightarrow Q\xi\alpha)$ equivale lógicamente a $(Q\xi\alpha \leftrightarrow \beta)$, también son válidas las fórmulas que se obtienen intercambiando $Q\xi\alpha$ y $\beta$ en la subfórmula que precede inmediatamente al segundo signo $\leftrightarrow$ en (15) y (16).

Siguiendo a Whitehead y Russell, llamaré *matriz* a una fórmula que no contiene cuantificadores. Una *fórmula prenexa* es una fórmula que consta de una fila de cuantificadores seguida de una matriz. La fila de cuantificadores

---

el cuantificador $Q_i$ en todas las posiciones en que dicho cuantificador la liga en $\psi$. Entonces, si $\psi$ es la formula $(\forall xP^1x_3 \rightarrow \forall x_2\exists x_1(P^2x_1x_2 \rightarrow \exists x_2\forall x_3(P^1x_2 \leftrightarrow P^2x_1x_3)))$, $\psi^*$ es la fórmula $(P^1x_3 \rightarrow \forall x_4\exists x_5(P^2x_5x_4 \rightarrow \exists x_6\forall x_7(P^1x_6 \leftrightarrow P^2x_5x_7)))$.

que precede a la matriz es el *prefijo* de la fórmula. Cada fórmula $\phi$ es lógicamente equivalente a una fórmula prenexa regular, que puede construirse a partir de $\phi$ en los cinco pasos siguientes:

(i) Se construye la fórmula regular $\phi^*$ correspondiente a $\phi$.

(ii) Se sustituye en $\phi^*$ cada subfórmula de la forma $(\alpha \rightarrow \beta)$ por otra de la forma $(\neg\alpha \vee \beta)$ y cada subfórmula de la forma $(\alpha \leftrightarrow \beta)$ por otra de la forma $((\alpha \wedge \beta) \vee (\neg\alpha \wedge \neg\beta))$; la fórmula $\phi_1$ obtenida mediante estas sustituciones es, como sabemos, lógicamente equivalente a $\phi^*$.

(iii) Se construye, mediante repetidas aplicaciones de (1), una fórmula $\phi_2$ lógicamente equivalente a $\phi_1$ en la cual no hay dos signos de negación consecutivos.

(iv) Mediante repetidas aplicaciones de (2)–(5), se obtiene una fórmula $\phi_3$ lógicamente equivalente a $\phi_2$ en que ningún signo de negación precede inmediatamente a un cuantificador.

(v) Mediante repetidas aplicaciones de (8)–(11), se obtiene una fórmula prenexa $\phi_4$ lógicamente equivalente a $\phi_3$.

Evidentemente, $(\phi \leftrightarrow \phi_4)$ es válida. Como se puede ver, gracias al paso (ii), no hace falta recurrir a (12)–(16).[42]

Una *fórmula prenexa de Skolem* es una fórmula prenexa en que cada cuantificador universal precede a todos los cuantificadores existenciales (SWL, p. 104).[43] Skolem (1920) estableció un importante resultado que en nuestros términos se deja enunciar así:

42 Observemos de paso que dos fórmulas prenexas son equivalentes si difieren sólo en el orden de los cuantificadores consecutivos de la misma clase. Pues es claro que $(\phi \leftrightarrow \psi)$ es válida, si $\phi$ es la fila $\alpha\forall\xi\forall\zeta\beta$ y $\psi$ es $\alpha\forall\zeta\forall\xi\beta$ o si $\phi$ es $\alpha\exists\xi\exists\zeta\beta$ y $\psi$ es $\alpha\exists\zeta\exists\xi\beta$, donde $\alpha$ es una fila de cuantificadores de longitud $\geq 0$ y $\beta$ es una fórmula prenexa o una matriz.

43 Muchos libros de texto —desde Hilbert y Ackermann 1928— definen, a la inversa, una fórmula prenexa de Skolem (o “fórmula en la forma normal prenexa de Skolem”) como una fórmula prenexa en que cada cuantificador existencial precede a todos los cuantificadores universales. Como ésta no es la definición de Skolem, llamo a las fórmulas que la satisfacen fórmulas prenexas de seudo-Skolem.

FPS *Si $\phi$ es una fórmula, hay una fórmula prenexa de Skolem $\phi^{\forall\exists}$ tal que $\phi$ es realizable si y sólo si $\phi^{\forall\exists}$ es realizable.*[44]

La demostración se facilita mediante el siguiente arreglo: si $\psi$ es una fórmula en la que $r$ variables diferentes $\xi_1, \ldots, \xi_r$ ocupan posiciones libres, introducimos un predicado $r$-ádico $\Pi$ (que no figura en $\psi$) que caracterizamos mediante la definición siguiente $\forall\xi_1 \ldots \forall\xi_r(\Pi\xi_1 \ldots \xi_r \leftrightarrow \psi)$. Es claro que $\psi$ equivale lógicamente a la fórmula elemental $\Pi\xi_1 \ldots \xi_r$.[45] Para aligerar el texto usaré la negrita para representar filas de símbolos similares, como se explica en la p. 491. Por ejemplo, $Q\boldsymbol{\zeta}^k$ representa una fila de $k$ cuantificadores mixtos. Sea $\phi$ una fórmula cualquiera y $\phi'$ una fórmula prenexa regular lógicamente equivalente a $\phi$ construida según el método descrito arriba. Si $\phi'$ no es una fórmula prenexa de Skolem, su prefijo contiene por lo menos una fila de cuantificadores existenciales consecutivos seguida inmediatamente por un cuantificador universal. Sea $\mathfrak{h}(\phi')$ el número de tales filas. Consideraremos en primer lugar el caso en que $\mathfrak{h}(\phi') = 1$. Entonces $\phi'$ tiene la forma $\boldsymbol{\forall\eta}^m\boldsymbol{\exists\xi}^h\boldsymbol{\forall\zeta}^k\alpha$, donde $\alpha$ es una matriz o una fórmula prenexa cuyo prefijo contiene sólo cuantificadores existenciales. Como $\alpha$ es regular, contiene $m + h + k + r$ variables libres diferentes ($r \geq 0$). Por lo tanto, $\phi'$ equivale lógicamente a

(17) $$\boldsymbol{\forall\eta}^m\boldsymbol{\exists\xi}^h\boldsymbol{\forall\zeta}^k\Pi\boldsymbol{\eta}^m\boldsymbol{\xi}^h\boldsymbol{\zeta}^k\boldsymbol{\chi}^r$$

donde $\Pi$ es un predicado $(m+h+k+r)$-ádico definido —como propuse arri-

44 Como la negación de una fórmula prenexa de Skolem equivale lógicamente —en virtud de (4) y (5)— a una fórmula prenexa de seudo-Skolem, el resultado enunciado equivale a este otro: *Si $\phi$ es una fórmula, hay una fórmula prenexa de seudo-Skolem $\phi^{\exists\forall}$ tal que $\phi$ es válida si y sólo si $\phi^{\exists\forall}$ es válida.*

45 La aplicación de un arreglo como éste a cada fórmula con variables libres —sugerida por la misma notación en el cálculo de relativos de Schröder empleado por Skolem— es legítima también en la versión del **CP1=** con un repertorio ilimitado de predicados que consideramos en este apéndice. En el caso de una fórmula $\phi$ perteneciente a una versión del **CP1=** sin tal repertorio ilimitado de predicados, se puede probar que hay una *extensión* de esta versión que se distingue de ella sólo en cuanto contiene cierto número adicional de predicados y en la cual puede construirse una fórmula prenexa de Skolem que es realizable en una interpretación de la versión extendida si y sólo si $\phi$ es realizable en una interpretación de la versión original.

ba— por $\forall\boldsymbol{\eta}^m\forall\boldsymbol{\xi}^h\forall\boldsymbol{\zeta}^k\forall\boldsymbol{\chi}^r(\Pi\boldsymbol{\eta}^m\boldsymbol{\xi}^h\boldsymbol{\zeta}^k\boldsymbol{\chi}^r \leftrightarrow \alpha)$. Sea $\Omega$ un predicado $(m+h)$-ádico que no figura en $\phi$. Postulamos que

(18) $$\forall\boldsymbol{\sigma}^m\forall\boldsymbol{\tau}^h(\Omega\boldsymbol{\sigma}^m\boldsymbol{\tau}^h \leftrightarrow \forall\boldsymbol{\chi}^r\forall\boldsymbol{\zeta}^k\Pi\boldsymbol{\sigma}^m\boldsymbol{\tau}^h\boldsymbol{\zeta}^k\boldsymbol{\chi}^r)$$

(donde la subfórmula que sigue inmediatamente a la doble flecha contiene precisamente $m + h$ variables libres). Ahora bien, (18) equivale lógicamente a

(19) $$\forall\boldsymbol{\sigma}^m\forall\boldsymbol{\tau}^h((\neg\Omega\boldsymbol{\sigma}^m\boldsymbol{\tau}^h \vee \forall\boldsymbol{\chi}^r\forall\boldsymbol{\zeta}^k\Pi\boldsymbol{\sigma}^m\boldsymbol{\tau}^h\boldsymbol{\zeta}^k\boldsymbol{\chi}^r) \wedge (\exists\boldsymbol{\chi}^r\exists\boldsymbol{\zeta}^k\neg\Pi\boldsymbol{\sigma}^m\boldsymbol{\tau}^h\boldsymbol{\zeta}^k\boldsymbol{\chi}^r \vee \Omega\boldsymbol{\sigma}^m\boldsymbol{\tau}^h))$$

y por ende a

(20) $$\forall\boldsymbol{\sigma}^m\forall\boldsymbol{\tau}^h\forall\boldsymbol{\chi}^r\forall\boldsymbol{\zeta}^k\exists\boldsymbol{\lambda}^r\exists\boldsymbol{\upsilon}^k((\neg\Omega\boldsymbol{\sigma}^m\boldsymbol{\tau}^h \vee \Pi\boldsymbol{\sigma}^m\boldsymbol{\tau}^h\boldsymbol{\zeta}^k\boldsymbol{\chi}^r) \wedge (\Omega\boldsymbol{\sigma}^m\boldsymbol{\tau}^h \vee \neg\Pi\boldsymbol{\sigma}^m\boldsymbol{\tau}^h\boldsymbol{\lambda}^k\boldsymbol{\upsilon}^r))$$

Por otra parte, en virtud de (18) y (7), tenemos que $\phi'$ equivale lógicamente a

(21) $$\forall\boldsymbol{\eta}^m\exists\boldsymbol{\xi}^h\Omega\boldsymbol{\eta}^m\boldsymbol{\tau}^h$$

La conjunción de (20) y (21) equivale a la siguiente fórmula prenexa de Skolem:

(22) $$\forall\boldsymbol{\eta}^m\forall\boldsymbol{\sigma}^m\forall\boldsymbol{\tau}^h\forall\boldsymbol{\chi}^r\forall\boldsymbol{\zeta}^k\exists\boldsymbol{\lambda}^r\exists\boldsymbol{\upsilon}^k\exists\boldsymbol{\xi}^h(\Omega\boldsymbol{\eta}^m\boldsymbol{\xi}^h \wedge ((\neg\Omega\boldsymbol{\sigma}^m\boldsymbol{\tau}^h \vee \Pi\boldsymbol{\sigma}^m\boldsymbol{\tau}^h\boldsymbol{\zeta}^k\boldsymbol{\chi}^r) \wedge (\Omega\boldsymbol{\sigma}^m\boldsymbol{\tau}^h \vee \neg\Pi\boldsymbol{\sigma}^m\boldsymbol{\tau}^h\boldsymbol{\lambda}^k\boldsymbol{\upsilon}^r)))$$

Es claro que (22) es realizable sólo si lo son (21) y su equivalente $\phi'$. Por otra parte, si $\phi'$ es realizable, esto es, si hay una interpretación $\langle\mathscr{D},f\rangle$ tal que $f(\phi') = 0$, hay también una interpretación $\langle\mathscr{D},f_1\rangle$ tal que $f_1$ concuerda con $f$ en todas las subfórmulas de $\phi'$ y $f_1(\Omega)$ es precisamente la clase de $(m+h)$-tuplos de elementos de $\mathscr{D}$ que, antepuestos a cualquier $(k+r)$-tuplo de tales elementos, integran la clase de $(m+h+k+r)$-tuplos $f_1(\Pi)$.[46] De esto se sigue que $f_1(\phi') = f_1((18)) = f_1((22)) = 0$ —ya que (22) equivale lógicamente a la

[46] En jerga matemática, $f_1(\Omega)$ es la proyección de $f_1(\Pi) \subseteq \mathscr{D}^{m+h} \times \mathscr{D}^{k+r}$ sobre $\mathscr{D}^{m+h}$.

conjunción de $\phi'$ y (18). Por lo tanto, (22) es realizable si $\phi'$ lo es. Con esto se completa la demostración de FPS en el caso especial en que $\mathfrak{h}(\phi') = 1$. En el caso general, $\phi'$ tiene la forma $Q\boldsymbol{\eta}^m\boldsymbol{\exists\xi}^h\boldsymbol{\forall\zeta}^k\alpha$, donde $Q\boldsymbol{\eta}^m$ es una fila mixta de cuantificadores que termina con un cuantificador universal. Mediante un razonamiento igual al anterior probamos que $\phi'$ es realizable si y sólo si lo es una fórmula $\psi$ que difiere de (22) únicamente en que $\psi$ comienza con la fila mixta $Q\boldsymbol{\eta}^m$ en vez de $\boldsymbol{\forall\eta}^m$. $\psi$ no es prenexa de Skolem, pero está más cerca de serlo que $\phi'$, por cuanto $\mathfrak{h}(\phi') = \mathfrak{h}(\psi) + 1$. Nuestro razonamiento es aplicable a $\psi$ y repitiéndolo un número finito de veces se obtiene una fórmula prenexa de Skolem que es realizable si y sólo si $\phi'$ es realizable. Dicho de otro modo: el resultado de Skolem que llamé FPS se establece para cualquier fórmula $\phi$ por inducción sobre el número $\mathfrak{h}$ de filas de cuantificadores existenciales consecutivos que preceden inmediatamente a un cuantificador universal en una fórmula prenexa ordinaria que equivalga lógicamente a $\phi$.

## XIV

## EL CÁLCULO DE PREDICADOS MONÁDICOS ES DECIDIBLE

Diré que un cálculo lógico es *decidible con respecto a la validez* o *v-decidible* si hay un algoritmo para determinar mediante un número finito de operaciones simples si una fórmula cualquiera de ese cálculo es o no válida. Como se narra en el Capítulo 2.7, Löwenheim (1915) demostró un teorema del cual se deduce que el cálculo de predicados monádicos —esto es, aquel fragmento del CP1= (o de otro cálculo lógico equivalente o análogo) que no contiene predicados poliádicos— es *v*-decidible. Bernays y Schönfinkel (1928, pp. 352–355) dieron una demostración muy elegante del mismo resultado, que presento aquí en una forma adaptada a la notación del Apéndice IX. Para no distraer al lector con detalles sutiles pero inesenciales, expondré primero una demostración aplicable a fórmulas sin functores, y luego daré las indicaciones necesarias para generalizarla.

Llamemos CP1$\boldsymbol{\mu}$ al fragmento del CP1 que sólo contiene predicados mo-

nádicos y 0-ádicos. Sea $\alpha$ una fórmula cualquiera del CP1**μ**. Traducido a nuestro modo de hablar, el teorema demostrado por Löwenheim dice que si $\alpha$ es verdadera en toda interpretación del CP1**μ** en un dominio finito (esto es, si no hay un entero positivo $n$ y una interpretación $\langle \mathcal{D}, f\rangle$ tal que $|\mathcal{D}| = n$ y $f(\alpha) = 1$), entonces $\alpha$ es válida (esto es, $f(\alpha) = 0$ cualquiera que sea la interpretación $\langle \mathcal{D}, f\rangle$). Ahora bien, si $\alpha$ no es válida, $\neg\alpha$ es realizable, de modo que la aseveración precedente equivale a esta otra: si $\beta$ es una fórmula del CP1**μ** y $\beta$ es realizable, entonces $\beta$ es realizable en un dominio finito (esto es, hay un entero positivo $n$ y una interpretación $\langle \mathcal{D}, f\rangle$ tal que $|\mathcal{D}| = n$ y $f(\beta) = 0$).

Sea $\phi$ una fórmula realizable del CP1**μ** en la que figuran $k$ predicados monádicos diferentes, $\Pi_1, \ldots, \Pi_k$; $\phi$ puede contener cualquier número finito $n \geq 0$ de predicados 0-ádicos ("variables proposicionales"). Demostraré que $\phi$ es realizable en un dominio $\mathcal{D}$ tal que $|\mathcal{D}| = 2^k$. Sea $\langle \mathcal{D}^*, f^*\rangle$ una interpretación tal que $|\mathcal{D}^*| > 2^k$ y $f^*(\phi) = 0$. Como sabemos, $f^*$ asigna a cada predicado $\Pi_h$ $(1 \leq h \leq k)$ un conjunto —posiblemente vacío— $f^*(\Pi_h) \subseteq \mathcal{D}^*$. Si $a$ es un elemento cualquiera de $\mathcal{D}^*$, $f^*$ le asocia lo que llamaré su *signatura* $\sigma(a)$ y que defino así: $\sigma(a)$ es un $k$-tuplo de ceros y unos cuyo $h$-ésimo término es igual a 0 si $a \in f^*(\Pi_h)$ y es igual a 1 si $a \notin f^*(\Pi_h)$. Obviamente, no hay más que $2^k$ signaturas diferentes entre las cuales se reparten todos los objetos del dominio $\mathcal{D}^*$, cualquiera que sea su cardinalidad. Ordenémoslas lexicográficamente y llamemos $\mathcal{D}_j$ al conjunto de elementos de $\mathcal{D}^*$ que comparten la $j$-ésima signatura. Sea $\mathcal{D} = \{\mathcal{D}_1, \ldots, \mathcal{D}_{2^k}\}$. Diré que una interpretación $\langle \mathcal{D}, f\rangle$ *se ajusta a* $\langle \mathcal{D}^*, f^*\rangle$ si satisface las tres condiciones siguientes: (i) $\mathcal{D}_j \in f(\Pi_h)$ si y sólo si $\mathcal{D}_j \subseteq f^*(\Pi_h)$ $(1 \leq h \leq k;\ 1 \leq j \leq 2^k)$; (ii) si $\xi$ es una variable individual que figura en $\phi$, $f(\xi) = \{x \in \mathcal{D}^* | \sigma(x) = \sigma(f^*(\xi))\}$,[47] y (iii) si $\Psi$ es un predicado 0-ádico que figura en $\phi$, $f(\Psi) = f^*(\Psi)$. Demostraré que si $\langle \mathcal{D}, f\rangle$ se ajusta a $\langle \mathcal{D}^*, f^*\rangle$ y, como hemos supuesto, $f^*(\phi) = 0$, también $f(\phi) = 0$. Ello es evidente si $\phi$ es una fórmula elemental $\Psi$ o $\Pi\xi$, y se establece fácilmente por inducción con respecto al número de conectivos si $\phi$ no contiene ningún cuantificador. Supongamos ahora que $\phi$ es una fórmula prenexa (Apéndice XIII). Razonamos por inducción con respecto al número $q(\phi)$ de cuantificadores en $\phi$. Ya vimos que la tesis puede demos-

[47] En otras palabras, $f(\xi) \in \mathcal{D}$ es el conjunto de objetos que comparten la signatura de $f^*(\xi) \in \Delta$.

trarse para el caso $q(\phi) = 0$. Supongamos que se la ha establecido también para fórmulas prenexas con $r$ cuantificadores, y que $q(\phi) = r + 1$. Entonces hay una variable $\xi$ y una fórmula prenexa $\psi$ con $r$ cuantificadores y los $k$ predicados $\Pi_1, \ldots ,\Pi_k$ tal que $\phi = \exists\xi\psi$ o $\phi = \forall\xi\psi$. En el primer caso, $f^*(\phi) = 0$ implica que $f_\xi{}^*(\psi) = 0$ en alguna $\xi$-variante $f_\xi{}^*$ de $\langle\mathcal{D}^*,f^*\rangle$ de suerte que, por la hipótesis inductiva, hay una $\xi$-variante $f_\xi$ de $\langle\mathcal{D},f\rangle$, ajustada a $\langle\mathcal{D}^*,f^*\rangle$, tal que $f_\xi(\psi) = 0$; por lo tanto, $f(\phi) = 0$. En el segundo caso, $f^*(\phi) = 0$ implica que $f_\xi{}^*(\psi) = 0$ en toda $\xi$-variante $f_\xi{}^*$ de $\langle\mathcal{D}^*,f^*\rangle$. Debido a la finitud de $\mathcal{D}$, cada $\xi$-variante $f_\xi$ de $\langle\mathcal{D},f\rangle$ se ajusta a alguna $\xi$-variante de $\langle\mathcal{D}^*,f^*\rangle$, de suerte que, por la hipótesis inductiva, $f_\xi(\psi) = 0$. Esto implica que $f(\phi) = 0$. Ahora bien, como se vio en la p. 519, si $\phi$ no es una fórmula prenexa, existe siempre una fórmula prenexa $\phi'$ lógicamente equivalente a $\phi$, cuyo valor, por ende, es idéntico al de $\phi$ en cualquier interpretación. Por lo tanto, $f^*(\phi) = 0$ implica que $f^*(\phi') = 0$; pero entonces, según acabamos de probar, $f(\phi') = 0$, de modo que $f(\phi) = 0$. Por lo tanto, si una fórmula cualquiera $\phi$ del CP1$\boldsymbol{\mu}$ es realizable en un dominio cualquiera, $\phi$ es realizable en un dominio finito.

De esto se sigue que el CP1$\boldsymbol{\mu}$ es $v$-decidible, ya que la realizabilidad o no realizabilidad de una fórmula dada $\phi$ en un dominio finito $\mathcal{D}$ puede determinarse en un número finito de pasos, calculando el valor de $\phi$ en cada una de las interpretaciones en $\mathcal{D}$ que $\phi$ admite.

Doy para terminar el complemento anunciado, que permite extender el resultado anterior a las fórmulas con functores. Llamemos CP1$\boldsymbol{\mu}$= al fragmento del CP1= que no contiene predicados poliádicos y supongamos que el argumento anterior se refiere a una fórmula $\phi$ del CP1$\boldsymbol{\mu}$=. Para que que el argumento sea concluyente tenemos que agregar una cuarta condición a las tres arriba prescritas para una interpretación $\langle\mathcal{D},f\rangle$ que se ajusta a $\langle\mathcal{D}^*,f^*\rangle$. La nueva condición concierne a la interpretación de los functores. Considérese la aplicación $\theta$: $\mathcal{D}^* \to \mathcal{D}$ que asigna a cada objeto $a$ el conjunto de los objetos que tienen la misma signatura que $a$. Como es habitual, llamo $\theta^n$ a la "aplicación producto" que envía cada $n$-tuplo $\langle a_1, \ldots ,a_n\rangle \in \mathcal{D}^{*n}$ al $n$-tuplo $\langle\theta(a_1), \ldots ,\theta(a_n)\rangle \in \mathcal{D}^n$. Entonces, diremos que $\langle\mathcal{D},f\rangle$ que se ajusta a $\langle\mathcal{D}^*,f^*\rangle$ sólo si, además de (i)–(iii), cumple la condición siguiente: (iv) si $\mathfrak{h}$ es un functor $n$-ario que figura en $\phi$, $f(\mathfrak{h}) \circ \theta^n = \theta \circ f^*(\mathfrak{h})$ (en otras palabras: si $f^*(\mathfrak{h})$ envía el $n$-tuplo $\langle a_1, \ldots ,a_n\rangle \in \mathcal{D}^{*n}$ a cierto objeto $a \in \mathcal{D}^*$, $f(\mathfrak{h})$ envía el $n$-tuplo $\langle\theta(a_1), \ldots ,\theta(a_n)\rangle \in \mathcal{D}^n$ precisamente a $\theta(a)$).

XV

# EL CÁLCULO PROPOSICIONAL ES COMPLETO

Como dije en el Capítulo 2.7, Post (1921) demostró que toda fórmula del cálculo proposicional que sea "positiva" en su interpretación algebraica —y por ende válida en la interpretación lógica habitual— es deducible por sustitución y *modus ponens* de los axiomas de su Postulado IV. La demostración es constructiva, en cuanto enseña a construir una deducción apropiada para cada fórmula positiva dada. Imitando a Post, la divido en cuatro etapas. Aunque la definición de fórmula que Post da en su Postulado I supone que los únicos conectivos son los signos de negación y disyunción, la demostración discurre como si los signos de conjunción, implicación y equivalencia también pertenecieran al cálculo.

[A] Digamos que una fórmula $\alpha$ tiene rango 0 —abreviado: $\Re(\alpha) = 0$— si consta únicamente de una variable proposicional sin conectivos; que $\Re(\neg\alpha) = 1 + \Re(\alpha)$, y que $\Re(\alpha \vee \beta) = 1 + \max(\Re(\alpha),\Re(\beta))$. Designemos con $\varphi(p)$ y $\varphi(q)$ a dos fórmulas cualesquiera tales que la segunda se deriva de la primera reemplazando uniformemente cierta variable proposicional $p$ por una variable proposicional $q$. Post prueba, por inducción sobre el rango de $\varphi(p)$, que toda aseveración de la forma $\vdash(p \leftrightarrow q) \rightarrow (\varphi(p) \leftrightarrow \varphi(q))$ es deducible de los postulados. Si $\Re(\varphi) = 0$, la tesis se reduce a una de las dos siguientes: (i) a $\vdash(p \leftrightarrow q) \rightarrow (p \leftrightarrow q)$, que se deriva por sustitución del familiar teorema $\vdash(p \rightarrow p)$;(ii) a $\vdash(p \leftrightarrow q) \rightarrow (r \leftrightarrow r)$, que se deriva por *modus ponens* y sustitución de $\vdash p \rightarrow (q \rightarrow p)$ y $\vdash r \leftrightarrow r$. Si la tesis se supone válida para fórmulas de rango menor que $m$, vale también si $\Re(\varphi) = m + 1$. En efecto, en tal caso $\varphi(p)$ puede escribirse en una de las formas $(\varphi_1(p) \vee \varphi_2(p))$ o $\neg\varphi_1(p)$, donde $\max(\Re(\varphi_1),\Re(\varphi_2)) = m$; y la tesis se deriva por *modus ponens* y sustitución de los teoremas $\vdash(p \leftrightarrow q) \rightarrow (p \leftrightarrow q)$, $\vdash(p \leftrightarrow q) \rightarrow (p \leftrightarrow q)$ y $\vdash(p \leftrightarrow q) \rightarrow (p \leftrightarrow q)$. Como Post señala en una nota, todos los asertos que aquí se invocan han sido deducidos en *Principia Mathematica*.

[B] Si $\varphi(p_1, \ldots ,p_k)$ es una fórmula que contiene $k$ variables proposicionales diferentes, se puede probar que hay una fórmula $\varphi'(p_1, \ldots ,p_k)$ tal que $\vdash\varphi(p_1, \ldots ,p_k) \leftrightarrow \varphi'(p_1, \ldots ,p_k)$, en la cual no hay otros conectivos binarios que $\vee$ y $\wedge$, y el signo $\neg$ sólo figura —si acaso— inmediatamente a la iz-

quierda de variables proposicionales. La prueba utiliza el metateorema demostrado bajo la letra [A] y tres teoremas deducidos en *Principia*, a saber, $\vdash\neg(p \vee q) \leftrightarrow (\neg p \wedge \neg q)$, $\vdash\neg\neg p \leftrightarrow p$, y $\vdash(p \leftrightarrow q) \rightarrow ((q \leftrightarrow r) \leftrightarrow (p \leftrightarrow r))$. [C] Invocando las propiedades distributivas y asociativas de $\vee$ y $\wedge$ se puede establecer que $\vdash\varphi'(p_1, \ldots, p_k) \leftrightarrow (\psi_1 \vee \ldots \vee \psi_n)$, donde cada subfórmula $\psi_i$ $(1 \leq i \leq n)$ es una conjunción de variables proposicionales o sus negaciones en la cual cada variable $p_j$ $(1 \leq j \leq k)$ figura a lo sumo una vez precedida por el signo $\neg$ y a lo sumo una vez sin ir precedida por dicho signo. Si la variable $p_k$ no figura en cierta subfórmula $\psi_i$, sustituimos esa subfórmula por $(\psi_i \wedge p_n) \vee (\psi_i \wedge \neg p_n)$. Sea $\psi'$ el producto final de estas sustituciones; es claro que $\vdash\varphi'(p_1, \ldots, p_k) \leftrightarrow \psi'$.[48] Apelando otra vez a las propiedades distributivas y asociativas de $\vee$ y $\wedge$ se construye una fórmula $\psi$ tal que:

(i) $$\vdash\psi' \leftrightarrow \psi$$

(ii) $$\psi = (\varphi_1 \wedge p_k \wedge \neg p_k) \vee (\varphi_2 \wedge p_k) \vee (\varphi_3 \wedge \neg p_k)$$

(iii) $p_k$ no figura en $\varphi_1$, $\varphi_2$ y $\varphi_3$

[D] Es claro entonces que $\vdash\varphi(p_1, \ldots, p_k) \leftrightarrow \psi$. Por lo tanto, si $\varphi(p_1, \ldots, p_k)$ es positiva, también $\psi$ es positiva (puesto que la positividad de una fórmula es una condición necesaria de su deducibilidad, como se mostró en las pp. 256-57). La deducibilidad de $\psi$ se prueba por inducción sobre el número de variables proposicionales que contiene. Si no contiene más que una, entonces $\psi$ es idéntica a la fórmula $(p \vee \neg p)$ o a la fórmula $((p \wedge \neg p) \vee p \vee \neg p)$ o se deduce de una de ellas por simple sustitución de una variable por otra. Ambas fórmulas son deducibles de los axiomas.[49] Supongamos ahora que $\psi$ contiene $k$ variables y la tesis se ha probado para cualquier fórmula positiva con menos de $k$ variables. Si $\psi$ es positiva, tanto $\varphi_2$ como $\varphi_3$ tienen que ser positivas.[50] Por la hipótesis inductiva, $\vdash\varphi_2$ y $\vdash\varphi_3$. De esto se deduce $\vdash\psi$,

48 De $\vdash p \vee \neg p$ y $\vdash p \rightarrow (q \leftrightarrow (p \wedge q))$ se deducen los teoremas $\vdash\psi_i \leftrightarrow (\psi_i \wedge (p_k \vee \neg p_k))$ y $\vdash(\psi_i \wedge (p_k \vee \neg p_k)) \leftrightarrow ((\psi_i \wedge p_k) \vee (\psi_i \wedge \neg p_k))$. Por lo tanto, el metateorema [A] autoriza a deducir $\vdash(\psi_1 \vee \ldots \vee \psi_n) \leftrightarrow \psi'$.

49 La segunda se deduce de la primera y el axioma $\vdash p \rightarrow (q \vee p)$.

50 Supongamos $\varphi_2$ no es positiva, de modo que $f(\varphi_2) = 1$ para cierta asignación $f$ de valores 0 ó 1 a las variables contenidas en $\varphi_2$. Extendamos $f$ a $p_k$ mediante la estipu-

aduciendo $\vdash p \to (q \to (p \leftrightarrow q))$, $\vdash((p \wedge q) \vee (p \wedge \neg q)) \leftrightarrow (p \wedge (q \vee \neg q))$, $\vdash p \to (q \to (p \wedge q))$ y $\vdash p \to (q \vee p)$. Por lo tanto, se deduce asimismo $\vdash \varphi(p_1, \ldots, p_k)$. Q.E.D.

## XVI

## UNA FORMA ABSTRACTA DEL PRIMER TEOREMA DE INCOMPLETUD DE GÖDEL (SMULLYAN 1992)

Como se explica en la Sección 2.10.2, Gödel 1931 demostró sus teoremas de incompletud para un cálculo bien determinado *P*, aunque haciendo hincapié en que la conclusión podía extenderse a una vasta familia de "sistemas afines". La siguiente versión "abstracta" del primer teorema, publicada por Smullyan (1992), aclara y precisa la índole de esa familia.

Para motivar la exposición utilizaré palabras como 'cálculo', 'oración', 'demostrable', 'verdadera', pero no hay que perder de vista que un "cálculo" aquí es un *conjunto cualquiera* y que las "oraciones" que llamo "verdaderas" o "demostrables" son sencillamente ciertas partes de ese conjunto que cumplen los requisitos que se indican. Por otra parte, cuando hablo de números me refiero a los números naturales (0, 1, 2, . . . ) en el sentido ordinario del término. Como se verá, la conclusión a que llegaremos sobre los "cálculos" depende solamente de ciertas relaciones de inclusión y exclusión entre sus partes y de dos relaciones, simples pero bien determinadas, entre el cálculo bajo consideración y el sistema $\mathbb{N}$ de los números.

Un conjunto no vacío $\mathscr{C}$ es un *cálculo* en la presente acepción si y sólo si:

---

lación $f(p_k) = 0$. Es claro entonces que $f(\psi) = 1$, lo cual contradice la suposición de que $\psi$ es positiva. Asimismo, si $\varphi_3$ no es positiva, tendremos que $f'(\varphi_3) = 1$ para cierta asignación $f'$ de valores 0 ó 1 a las variables contenidas en $\varphi_3$. Extendamos $f'$ a $p_k$ mediante la estipulación $f'(p_k) = 1$. Es claro entonces que $f'(\psi) = 1$, lo cual contradice la suposición de que $\psi$ es positiva.

(I) $\mathscr{C}$ incluye (i) un conjunto denumerable $\mathscr{E}$ de *expresiones*; (ii) un conjunto $\mathbb{O} \subseteq \mathscr{E}$ de *oraciones*; (iii) un conjunto $\mathscr{D} \subseteq \mathbb{O}$ de *oraciones demostrables*; (iv) un conjunto $\mathscr{R} \subseteq \mathbb{O}$ de *oraciones refutables*; (v) un conjunto $\mathscr{V} \subseteq \mathbb{O}$ de *oraciones verdaderas*; (vi) un conjunto $\mathscr{H} \subseteq \mathscr{E}$ de *predicados*;

(II) existe una aplicación $\Phi: \mathscr{E} \times \mathbb{N} \rightarrow \mathscr{E}$ tal que $\Phi(\mathscr{H} \times \mathbb{N}) \subseteq \mathbb{O}$ (la imagen de $\mathscr{H} \times \mathbb{N}$ por $\Phi$ está contenida en $\mathbb{O}$).

La condición (I) se entiende sin dificultad, si uno se abstiene de atribuir un significado específico a las palabras en cursiva, introducidas únicamente para motivar la exposición. La aplicación postulada en (II) asigna a cada expresión $E \in \mathscr{E}$ y a cada número $n$ una expresión única que llamaré $E(n)$. En particular, si la expresión es un predicado $H$, $H(n)$ es una oración. Si $H(n) \in \mathscr{V}$, decimos que el número $n$ satisface el predicado $H$ y que $H$ se aplica a $n$. Decimos que el predicado $H$ *expresa* el conjunto numérico $K$ si $K$ es precisamente el conjunto de los números a que $H$ se aplica. En otras palabras, $H$ expresa $K$ si y sólo si $H(n) \in \mathscr{V} \leftrightarrow n \in K$. El conjunto $K$ es *expresable* en $\mathscr{C}$ si hay un $H \in \mathscr{H} \subseteq \mathscr{C}$ que expresa a $K$.

$\mathscr{C}$ es un cálculo *correcto* si toda oración demostrable es verdadera y no lo es ninguna oración refutable. Nos interesa establecer las condiciones suficientes para que un cálculo correcto contenga una oración verdadera que no es demostrable. (Esto es, la condición suficiente para que $\mathscr{V} \backslash \mathscr{D} \neq \varnothing$, dado que $\mathscr{D} \subseteq \mathscr{V}$ y $\mathscr{R} \cap \mathscr{V} = \varnothing$).

Como $\mathscr{E}$ es numerable, podemos asignar un número de identidad $\#[E]$ a cada expresión $E$. Supondremos que la aplicación $\#:\mathscr{E} \rightarrow \mathbb{N}$ es biyectiva, de modo que cada número natural es el número de identidad de alguna expresión. (véase el Apéndice XVII). Llamaré $E_n$ a la expresión cuyo número de identidad es $n$. En otras palabras, $n = \#[E_n]$. ($E_n = \dagger[n]$ en la nomenclatura del Capítulo 2.10). La *diagonalización* de $E_n$ es la expresión $E_n(n)$. Si $E_n$ es un predicado, $E_n(n)$ es una oración. Esta oración es verdadera si y sólo si el predicado $E_n$ se aplica a su propio número de identidad. Las estipulaciones anteriores definen completamente la función $\Delta: \mathbb{N} \rightarrow \mathbb{N}$ por $n \mapsto \#[E_n(n)]$. $\Delta$ es la *función diagonal* del sistema determinado por $\mathscr{C}$ y $\#$.

Si $K \subseteq \mathbb{N}$, llamo $K'$ al conjunto $\mathbb{N} \backslash K$, esto es, el complemento de $K$ con respecto a $\mathbb{N}$. Normalmente llamaríamos $\Delta^{-1}(K)$ a la imagen inversa de $K$ bajo la función diagonal $\Delta$, pero para aligerar la tipografía, la llamaré $K^*$

(como hace Smullyan). En otras palabras: $n \in K^* \leftrightarrow \Delta(n) \in K$.

Sea $D$ el conjunto de los números de identidad de las oraciones demostrables de $\mathcal{C}$: $D = \{x{:}x \in \mathbb{N} \wedge E_x \in \mathcal{D}\}$. Probaremos el siguiente

TEOREMA: Si $\mathcal{C}$ es un cálculo correcto y el conjunto $(D')^*$ es expresable en $\mathcal{C}$, $\mathcal{V} \setminus \mathcal{D} \neq \varnothing$.

DEMOSTRACIÓN: Supongamos que $\mathcal{C}$ es correcto y que el predicado $H$ expresa el conjunto $(D')^*$. Sea $h = \#[H]$. Comprobaremos que la oración $H(h)$ es verdadera pero no es demostrable. Si $H$ expresa $(D')^*$, entonces, para cada número $n$, $H(n) \in \mathcal{V} \leftrightarrow n \in (D')^*$. En particular, $H(h) \in \mathcal{V} \leftrightarrow h \in (D')^*$. Ahora bien, $h \in (D')^* \leftrightarrow \Delta(h) \in D' \leftrightarrow \Delta(h) \notin D$. Por definición, $\Delta(h) = \#[E_h(h)] = \#[H(h)]$ (puesto que $h = \#[H]$). Por lo tanto, $\Delta(h) \notin D \leftrightarrow H(h) \notin \mathcal{D}$. Uniendo nuestras tres cadenas de equivalencias, concluimos que $H(h) \in \mathcal{V} \leftrightarrow H(h) \notin \mathcal{D}$. Tenemos pues que, o bien $H(h) \in \mathcal{D} \setminus \mathcal{V}$, o bien $H(h) \in \mathcal{V} \setminus \mathcal{D}$. Si $\mathcal{C}$ es correcto, $\mathcal{D} \setminus \mathcal{V} = \varnothing$. Por lo tanto, $H(h) \in \mathcal{V} \setminus \mathcal{D} \neq \varnothing$.

## XVII

## NÚMEROS DE GÖDEL: UNA ALTERNATIVA

En las pp. 329-30 se explica el método adoptado por Gödel (1931) para asignarle un número de identidad —lo que he llamado un 'gödel'— a cada objeto de un cálculo lógico. El método de Gödel es muy elegante pero no permite determinar con facilidad cuál es el número asignado a cierto objeto, o el objeto que corresponde a cierto número. (De hecho, hay números a los que no corresponde ningún objeto). Desde 1931 se han introducido varios otros métodos. En las pp. 384-85 expliqué el adoptado por Turing para asignarle un gödel a sus programas de cómputo. Siguiendo a Smullyan (1992), aplicaré aquí una variante del método de Turing a los objetos de un cálculo lógico. Para fijar ideas, me refiero al cálculo predicativo de segundo orden, CP2=, descrito en el Apéndice IX.H, pero el método se puede adaptar fácil-

mente a cualquier otro cálculo.

Para empezar, se le asigna un dígito a cada signo del cálculo. El número denotado por ese dígito es el gödel de ese signo. El gödel de una fila de signos es el número denotado por la correspondiente fila de dígitos. Para que este método de identificación pueda extenderse a las listas de filas de signos, introducimos un nuevo signo ◇, para separar las filas de una lista. Estipulamos que (i) una fila de signos del cálculo que no contiene el signo ◇ es una *expresión*; (ii) si α es una expresión, α◇ es una *lista* de una expresión, y (iii) si α es una expresión y β es una lista de *n* expresiones, α◇β es una lista de *n*+1 expresiones (*n* ≥ 1). El CP2= tiene 18 signos, a saber, ¬, →, ∧, ∨, ↔, ∀, ∃, =, (, ), $_|$, *, *x*, *X*, *F*, *a*, *P*, *f*. Para aplicarle nuestro método de identificación tenemos, pues, que utilizar un sistema numérico con 19 dígitos, esto es, un sistema de base 19. Los programadores, que usan un sistema de base 16, designan los números del 0 al 9 con los dígitos corrientes y los números 10, 11, 12, 13, 14 y 15 con las letras A, B, C, D, E y F, respectivamente. Sigamos la misma convención, poniendo G = 16, H = 17 y J = 18 (omito la I porque puede confundirse con el 1). Si asignamos el 0 al signo ◇ y los restantes dígitos a los signos del CP2= en el orden en que los presenté arriba, tenemos que el gödel de la fórmula ∀*xPx* es 6DHD, que equivale, en nuestro sistema decimal, a $6\times19^3 + 13\times19^2 + 17\times19 + 13 =$ 46.183. Por otro lado, es claro que cualquier número es el gödel de un signo, una fila de signos o una lista de filas de signos. Por ejemplo, el número 1.000.000 = $7\times19^4 + 12\times19^3 + 15\times19^2 + 1\times19 + 11$ = 7CF1B es el gödel de la expresión ∃**X*¬$_|$, la cual, claro está, no es una fórmula. En situaciones especiales será oportuno adoptar otras convenciones. Por ejemplo, en el cálculo que usa Smullyan, diseñado para representar la aritmética, hay una constante 0 que denota el cero, y un functor ′, ‘el siguiente de’, que se escribe pospuesto a su argumento. Así la expresión 0′′′′ denota el número 4. Smullyan asigna el dígito 1 al signo 0 y el dígito 0 al signo ′. Con estas convenciones, el gödel de la expresión que denota el número *n* se representa muy convenientemente —en el sistema numérico utilizado por Smullyan— con un 1 seguido de *n* ceros.

# XVIII

## LOS AXIOMAS DEL CÁLCULO DE PRIMER ORDEN INVESTIGADO POR GÖDEL (1930) SON DERIVABLES EN EL CÁLCULO DE SECUENTES PROPUESTO POR GENTZEN (1938)

Doy aquí varios ejemplos de derivaciones en el cálculo de secuentes utilizado por Gentzen en su segunda demostración de la consistencia de la aritmética elemental (1938). Aunque el principal propósito de este apéndice es ilustrar el concepto de derivación en dicho cálculo, explicado en la Sección 2.12.1, aprovecho la oportunidad para mostrar que cada uno de los axiomas 1-6 del cálculo de primer orden que Gödel (1930) demostró completo (Capitulo 2.8) es la posfórmula (única) de un secuente sin prefórmulas derivable en el cálculo de Gentzen, y por lo tanto también se puede invocar como axioma en este último.[51]

Presento las derivaciones en forma de esquemas. Igual que en la presentación de los esquemas de inferencias en la Sección 2.12.1, las mayúsculas griegas representan listas de fórmulas separadas por comas; las mayúsculas góticas representan fórmulas y las minúsculas góticas, variables. A la derecha de cada conclusión anoto el tipo de inferencia utilizado para derivarla. Antes de derivar los axiomas de Gödel, doy el esquema de un "hilo" —en el sentido explicado en la Sección 2.12.2— que suele figurar en las derivaciones: partiendo de un secuente de la forma $\Gamma, \mathfrak{A} \twoheadrightarrow \Delta, \mathfrak{B}$ se puede derivar siempre el secuente correspondiente $\Gamma \twoheadrightarrow \Delta, \neg\mathfrak{A} \vee \mathfrak{B}$.

[51] Cada secuente ➛ A, en que A es uno de los seis axiomas en cuestión, puede insertarse como secuente inicial en cualquier derivación, sobreentendiéndose que sobre él va una derivación como las dadas aquí. Obsérvese que los axiomas 7 y **8** de Gödel no pueden justificarse así, porque contienen el signo =, el cual no es un símbolo lógico del cálculo de Gentzen. Pero el axioma **7** ($x = x$) es evidentemente la posfórmula única de un secuente básico matemático sin prefórmulas. También lo es el axioma **8**, si exigimos que las fórmulas $F(x)$ y $F(y)$ que figuran en él sean elementales.

$$\cfrac{\cfrac{\cfrac{\cfrac{\cfrac{\Gamma, \mathfrak{A} \rightarrow \Delta, \mathfrak{B}}{\Gamma \rightarrow \Delta, \mathfrak{B}, \neg\mathfrak{A}}\ (\neg)}{\Gamma \rightarrow \Delta, \mathfrak{B}, \neg\mathfrak{A} \vee \mathfrak{B}}\ (\vee)}{\Gamma \rightarrow \Delta, \neg\mathfrak{A} \vee \mathfrak{B}, \mathfrak{B}}\ (\text{Permutación})}{\Gamma \rightarrow \Delta, \neg\mathfrak{A} \vee \mathfrak{B}, \neg\mathfrak{A} \vee \mathfrak{B}}\ (\vee)}{\Gamma \rightarrow \Delta, \neg\mathfrak{A} \vee \mathfrak{B}}\ (\text{Contracción})$$

Omitiendo las líneas 2-5 (y las rayas horizontales que las preceden), el esquema precedente puede condensarse en el siguiente esquema de inferencia auxiliar que, siguiendo el ejemplo de los manuales de lógica, llamo *teorema de la deducción* (TD).

TD

$$\frac{\Gamma, \mathfrak{A} \rightarrow \Delta, \mathfrak{B}}{\Gamma \rightarrow \Delta, \neg\mathfrak{A} \vee \mathfrak{B}}$$

Ahora derivo los axiomas **1**-**6**. Para ahorrar espacio, no anoto las permutaciones efectuadas.

**1**

$$\cfrac{\cfrac{\mathfrak{A} \rightarrow \mathfrak{A} \qquad \mathfrak{A} \rightarrow \mathfrak{A}\ (\text{SB} \mid \text{SB})}{\mathfrak{A} \vee \mathfrak{A} \rightarrow \mathfrak{A}}\ (\vee)}{\rightarrow \neg(\mathfrak{A} \vee \mathfrak{A}) \vee \mathfrak{A}}\ (\text{TD})$$

**2**

$$\frac{\dfrac{\mathfrak{A} \rightarrow \mathfrak{A}}{\mathfrak{A} \rightarrow \mathfrak{A} \vee \mathfrak{B}}}{\rightarrow \neg\mathfrak{A} \vee (\mathfrak{A} \vee \mathfrak{B})}$$

(SB)
(∨)
(TD)

**3**

$$\frac{\dfrac{\dfrac{\mathfrak{A} \rightarrow \mathfrak{A}}{\mathfrak{A} \rightarrow \mathfrak{B} \vee \mathfrak{A}} \qquad \dfrac{\mathfrak{B} \rightarrow \mathfrak{B}}{\mathfrak{B} \rightarrow \mathfrak{B} \vee \mathfrak{A}}}{\mathfrak{A} \vee \mathfrak{B} \rightarrow \mathfrak{B} \vee \mathfrak{A}}}{\rightarrow \neg(\mathfrak{A} \vee \mathfrak{B}) \vee (\mathfrak{B} \vee \mathfrak{A})}$$

(SB | SB)
(∨|∨)
(∨)
(TD)

**4**

$$\frac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\mathfrak{C} \rightarrow \mathfrak{C}}{\neg\mathfrak{A}, \mathfrak{C} \rightarrow \mathfrak{C}} \qquad \dfrac{\dfrac{\mathfrak{A} \rightarrow \mathfrak{A}}{\mathfrak{A} \rightarrow \mathfrak{A}, \mathfrak{C}}}{\neg\mathfrak{A}, \mathfrak{A} \rightarrow \mathfrak{C}}}{\dfrac{\neg\mathfrak{A}, \mathfrak{C} \vee \mathfrak{A} \rightarrow \mathfrak{C}}{\neg\mathfrak{A}, \mathfrak{C} \vee \mathfrak{A} \rightarrow \mathfrak{C} \vee \mathfrak{B}}} \qquad \dfrac{\dfrac{\mathfrak{B} \rightarrow \mathfrak{B}}{\mathfrak{B} \rightarrow \mathfrak{C} \vee \mathfrak{B}}}{\mathfrak{C} \vee \mathfrak{A}, \mathfrak{B} \rightarrow \mathfrak{C} \vee \mathfrak{B}}}{\neg\mathfrak{A} \vee \mathfrak{B}, \mathfrak{C} \vee \mathfrak{A} \rightarrow \mathfrak{C} \vee \mathfrak{B}}}{\neg\mathfrak{A} \vee \mathfrak{B}, \rightarrow \neg(\mathfrak{C} \vee \mathfrak{A}) \vee (\mathfrak{C} \vee \mathfrak{B})}}{\rightarrow \neg(\neg\mathfrak{A} \vee \mathfrak{B}) \vee \neg(\mathfrak{C} \vee \mathfrak{A}) \vee (\mathfrak{C} \vee \mathfrak{B})}$$

(SB)
(SB|deb.)
(deb.| ¬|SB)
(∨|∨)
(∨|deb.)
(∨)
(TD)
(TD)

**5**

$$\mathfrak{F}(\mathfrak{y}/\mathfrak{x}) \rightarrow \mathfrak{F}(\mathfrak{y}/\mathfrak{x})$$

---

$$\forall \mathfrak{x}\mathfrak{F} \rightarrow \mathfrak{F}(\mathfrak{y}/\mathfrak{x}) \qquad (\forall)$$

---

$$\rightarrow \neg\forall \mathfrak{x}\mathfrak{F} \vee \mathfrak{F}(\mathfrak{y}/\mathfrak{x}) \qquad \text{(TD)}$$

**6**

$$\mathfrak{A} \rightarrow \mathfrak{A} \qquad\qquad \mathfrak{F}(\mathfrak{y}/\mathfrak{x}) \rightarrow \mathfrak{F}(\mathfrak{y}/\mathfrak{x}) \qquad \text{(SB|SB)}$$

---

$$\mathfrak{A} \rightarrow \mathfrak{A}, \forall \mathfrak{x}\mathfrak{F} \qquad\qquad \mathfrak{F}(\mathfrak{y}/\mathfrak{x}) \rightarrow \mathfrak{F}(\mathfrak{y}/\mathfrak{x}), \mathfrak{A} \qquad \text{(deb.|deb.)}$$

---

$$\mathfrak{A} \rightarrow \mathfrak{A}, \forall \mathfrak{x}\mathfrak{F}, \mathfrak{F}(\mathfrak{y}/\mathfrak{x}) \qquad \mathfrak{F}(\mathfrak{y}/\mathfrak{x}) \rightarrow \mathfrak{F}(\mathfrak{y}/\mathfrak{x}), \mathfrak{A}, \forall \mathfrak{x}\mathfrak{F} \qquad \text{(deb.|deb.)}$$

---

$$\mathfrak{A} \vee \mathfrak{F}(\mathfrak{y}/\mathfrak{x}) \rightarrow \mathfrak{A}, \forall \mathfrak{x}\mathfrak{F}, \mathfrak{F}(\mathfrak{y}/\mathfrak{x}) \qquad (\vee)$$

---

$$\mathfrak{A} \vee \mathfrak{F}(\mathfrak{y}/\mathfrak{x}) \rightarrow \mathfrak{A}, \forall \mathfrak{x}\mathfrak{F}, \forall \mathfrak{x}\mathfrak{F} \qquad (\forall)$$

---

$$\mathfrak{A} \vee \mathfrak{F}(\mathfrak{y}/\mathfrak{x}) \rightarrow \mathfrak{A}, \forall \mathfrak{x}\mathfrak{F} \qquad \text{(contr.)}$$

---

$$\mathfrak{A} \vee \mathfrak{F}(\mathfrak{y}/\mathfrak{x}) \rightarrow \mathfrak{A} \vee \forall \mathfrak{x}\mathfrak{F}, \forall \mathfrak{x}\mathfrak{F} \qquad (\vee)$$

---

$$\mathfrak{A} \vee \mathfrak{F}(\mathfrak{y}/\mathfrak{x}) \rightarrow \mathfrak{A} \vee \forall \mathfrak{x}\mathfrak{F}, \mathfrak{A} \vee \forall \mathfrak{x}\mathfrak{F} \qquad (\vee)$$

---

$$\mathfrak{A} \vee \mathfrak{F}(\mathfrak{y}/\mathfrak{x}) \rightarrow \mathfrak{A} \vee \forall \mathfrak{x}\mathfrak{F} \qquad \text{(contr.)}$$

---

$$\rightarrow \neg(\mathfrak{A} \vee \mathfrak{F}(\mathfrak{y}/\mathfrak{x})) \vee (\mathfrak{A} \vee \forall \mathfrak{x}\mathfrak{F}) \qquad \text{(TD)}$$

(donde $\mathfrak{y}$ es una variable que no figura en la fórmula $\mathfrak{A}$)

## XIX

## ALGUNAS IDEAS DE BROUWER

Luitzen Egbertus Jan Brouwer (1881-1966) alcanzó muy joven una concepción de la matemática radicalmente opuesta a la de los autores que estudiamos en este libro. Ella se perfila claramente en su tesis doctoral (1907), sobre todo en la primera versión, cuyos pasajes más audaces suprimió por recomendación de su consejero, D. J. Korteweg (van Stigt, 1979). Korteweg le sugirió a su joven y extraordinariamente brillante discípulo que se aplicara primero a resolver problemas de la matemática tradicional, para asegurar su autoridad y prestigio. En cumplimiento de este plan, Brouwer sentó las bases de la teoría topológica de la dimensión, demostrando que el número de dimensiones de un espacio es invariante bajo biyecciones bicontinuas (1911, 1913). Tras este importante logro, con el cual rescató una intuición que muchos creían destruida por Cantor (*vide* p. 26), Brouwer empezó a publicar artículos de carácter filosófico en las Actas de la Academia Neerlandesa de Ciencias y en revistas internacionales de matemáticas.

Brouwer solía decir que sus ideas sobre la matemática emanaban de su personal concepción del mundo y de la vida. Esta tiene sólo un tenue vínculo —a través de Schopenhauer— con la tradición filosófica europea, y me confieso incapaz de comprenderla y explicarla.[52] Por suerte, aquí busco sólo marcar el contraste entre Brouwer y el conjuntismo, y para ello no es preciso calar muy hondo. Me limitaré a citar algunas palabras de Brouwer sobre la intuición fundamental que según él es la raíz de las matemáticas, para luego, a partir de ahí, resumir sus ideas sobre el infinito matemático y el principio del tercero excluido.

Según Brouwer, el "fenómeno fundamental del intelecto humano" es "la disociación de los momentos de la vida en partes cualitativamente diversas, que sólo pueden reunirse en cuanto permanezcan separadas por el tiempo";

[52] Cf. especialmente sus artículos "Matemática, ciencia y lenguaje" (1929) y "Conciencia, filosofía y matemáticas" (1948), ambos reproducidos en Brouwer, CW, tomo I. En el mismo tomo, pp. 1-10, pueden consultarse extractos, traducidos al inglés, del libro juvenil de Brouwer, *Vida, arte y mística* (1905).

hecha abstracción de su "contenido emocional" este fenómeno pasa a ser "el fenómeno fundamental del pensamiento matemático, la intuición de la nuda duo-unidad (*two-oneness*)" (1913a; CW I, 127).

> Esta intuición de la duo-unidad, la intuición básica de las matemáticas, crea no sólo los números uno y dos, sino también todos los números ordinales finitos, en cuanto uno de los elementos de la duo-unidad puede pensarse como una nueva duo-unidad, y este proceso puede repetirse indefinidamente. Ello da nacimiento además al más pequeño número ordinal infinito ω. Por último, esta intuición básica de las matemáticas, en la que se unen lo conexo y lo separado, lo continuo y lo discreto, genera inmediatamente la intuición del continuo lineal, es decir, del "entre" que no se agota con la interposición de nuevas unidades y que, por lo tanto, no puede nunca pensarse como una mera colección de unidades.
>
> (Brouwer 1913a; CW I, 127-28)

En la ponencia que presentó al IV Congreso Internacional de Matemáticos, titulada "Las potencias posibles" (1908a),[53] Brouwer fue más explícito. En la intuición originaria de la duo-unidad (*Zweieinigkeit*) se reúnen las intuiciones de lo continuo y lo discreto, "justamente porque lo segundo no se piensa por sí mismo, sino bajo retención del recuerdo de lo primero" (CW I, 102). En este tener o sostener a lo uno *con* lo otro (*Zusammenhaltung*) consiste precisamente la intuición de lo *con*-tinuo. Vaciada de su contenido, esta "sensación" (*Empfindung)* de lo fijo y lo evanescente juntos, de lo permanente y lo cambiante juntos, es la intuición matemática originaria. En ella se basan los dos desarrollos siguientes:

(1) Uno piensa la intuición de la duo-unidad como una nueva unidad, a la que se agrega una nueva alteridad —llamada 'tres'— y así sucesivamente; de este modo se establece el tipo de orden de los números naturales, que Brouwer, como Cantor, llama ω.
(2) Al vivir la intuición originaria como *tránsito* entre lo primero y lo segundo se establece el tipo de orden η de los racionales mayores que 0 y menores que 1, ordenados de menor a mayor.

53 Recuérdese que Cantor llamó 'potencia' (*Mächtigkeit*) a lo que en este libro llamamos 'numerosidad'; cf. p. 21.

Según esto, no puede haber más que una numerosidad infinita, a saber, la denumerable. Con todo, Brouwer admite que se hable de una numerosidad superior en cualquiera de los dos sentidos siguientes:

(*a*) Cada conjunto denumerable dado perteneciente a un sistema matemático genera un nuevo elemento que también pertenece al sistema. Por esta vía sólo pueden construirse conjuntos denumerables, no el sistema completo, porque este no puede ser denumerable. "Es incorrecto considerar a este sistema entero como un conjunto matemático, pues no es posible terminar de edificarlo desde la intuición matemática originaria" (1908a; CW I, 103). Brouwer propone estos ejemplos: la totalidad de los números de la segunda clase (*vide* pp. 37-38), la totalidad de los puntos definibles del continuo, la totalidad de los sistemas matemáticos.

(*b*) Al continuo *entre* lo primero y lo segundo se lo puede considerar como una *matriz generadora* de unidades o puntos, y postularse que dos puntos deben considerarse distintos si y sólo si es posible distinguir sus respectivas posiciones en una cierta escala de tipo de orden $\eta$. 'Se observa entonces que el continuo definido de este modo no puede agotarse nunca como matriz de puntos" (1908a; CW I, 103).

Brouwer concluye que existe sólo una "potencia" o numerosidad para conjuntos matemáticos infinitos, "a saber, la denumerable". A ella cabe agregar (*a*) *la denumerablemente inconclusa*, "pero ello denota un *método*, no un conjunto"; y (*b*) *la continua*, la cual, sí, "denota algo terminado (*etwas Fertiges*), pero sólo como *matriz*, no como conjunto" (1908a; CW I, 104).

A Brouwer se lo conoce en los círculos filosóficos sobre todo porque negó la validez universal del principio lógico del tercero excluido. El alcance y la justificación de su rechazo sólo se puede apreciar contra el trasfondo de las ideas precedentes. Brouwer no estaba en el negocio de crear una lógica alternativa, como la que, revestida de una parafernalia formal muy semejante a la que detestaba en sus adversarios, circula con el nombre de "lógica intuicionista". Para él, la actividad matemática, alimentada de la intuición originaria, es extralingüística. El lenguaje matemático no es más que un recurso defectuoso de los hombres para comunicarse las matemáticas unos a otros y para reforzar su memoria de las matemáticas (Brouwer 1907; CW I, 92). La *verdad* sólo se encuentra "en la *realidad*, esto es, en las experiencias presentes y pasadas de la conciencia", las cuales incluyen cosas y sus cualidades, emociones, reglas (jurídicas, de cooperación, de juego), actos mate-

riales, actos de pensamiento, actos matemáticos (*mathematical deeds*).

> Las verdades suelen trasmitirse con palabras o complejos de palabras, generalmente tomados en préstamo de los lenguajes cooperativos, de tal modo que para el sujeto una cierta palabra o complejo de palabras siempre evoca una verdad determinada [ . . . ]. Hay además un sistema de reglas generales llamado *lógica*, que permite al sujeto deducir de sistemas de complejos de palabras que trasmiten verdades, otros complejos de palabras que generalmente también trasmiten verdades. [ . . . ] Ello no significa que dichos complejos adicionales de palabras trasmitan verdades *antes* de que estas verdades se hayan experimentado, ni que estas verdades *siempre puedan* experimentarse. En otras palabras, la lógica no es un instrumento fiable para descubrir verdades y no puede deducir verdades que no fueran accesibles también de otra manera.
>
> (Brouwer 1948; CW I, 488)

Brouwer se refiere expresamente a tres principios lógicos: el principio del *silogismo*,[54] el principio de *contradicción* y el principio *del tercero excluido*. Con respecto a ellos se plantea la cuestión siguiente:

> *Supóngase que una construcción matemática intuicionista se ha descrito cuidadosamente con palabras y que luego, ignorando de momento el carácter introspectivo de la construcción matemática, su descripción lingüística es considerada por sí misma y sometida a la aplicación lingüística de un principio de la lógica clásica. ¿Es siempre posible entonces ejecutar una construcción matemática sin palabras que halle expresión en la figura lógico-lingüística en cuestión?*
>
> (Brouwer 1952; CW I, 510; cursiva de Brouwer)

Haciendo reserva de “la inevitable inadecuación del lenguaje como modo de descripción”, Brouwer admite una respuesta *afirmativa* en el caso de los principios de contradicción y del silogismo. En cambio, “salvo en casos especiales”, la respuesta es *negativa* en el caso del principio del tercero excluido, de modo que “este último principio debe rechazarse como instrumento

[54] Este principio infiere “de la inclusión de un sistema *b* en un sistema *c* combinada con la inclusión de un sistema *a* en el sistema *b*, la inclusión directa de *a* en *c”* (Brouwer 1908b; CW I, 109). Se trata, en suma, del esquema del modo *barbara*. ¿Pensaba Brouwer que todo silogismo podría reducirse a este solo esquema, sin recurrir a otros principios (excepto quizás el de contradicción)?

para descubrir nuevas verdades matemáticas" (Ibid.). La validez de este principio, dice Brouwer, equivale a la tesis de que no hay problemas matemáticos insolubles, para sostener la cual no tenemos el más mínimo indicio (1908b; CW I, 109).

Naturalmente, el principio es enteramente admisible cuando se trata de sistemas finitos. Cada construcción de carácter finito y acotado en un sistema de esta clase puede intentarse solamente en un número finito de formas diferentes; cada intento, o bien tendrá éxito, o continuará hasta el punto en que se vuelva imposible. Por lo tanto, se puede siempre alcanzar una decisión sobre cualquier aserto concerniente a la posibilidad de una construcción de carácter finito y acotado en un sistema matemático finito.

Para mostrar que ello no es así en el caso de los sistemas infinitos, Brouwer introduce el concepto de una propiedad fugitiva de los números naturales (1929; CW I, p. 425). Diremos que la propiedad $f$ es *fugitiva* si satisface las condiciones siguientes:

1. Para cada número natural $n$ es posible decidir si $n$ tiene $f$ o si es absurdo que $n$ tenga $f$.
2. No se conoce ningún método para calcular un número $n$ que tenga $f$.
3. No se puede probar que sea absurda la hipótesis de que existe un número natural que tiene $f$.

He aquí un ejemplo sencillo. Digamos que el número natural $n$ tiene la propiedad $f_\pi$ si y sólo el $n$-ésimo, el ($n$+1)-ésimo, el ($n$+2)-ésimo,…, y el ($n$+9)-ésimo dígito en la expansión decimal de $\pi$, tomados en ese orden, forman la secuencia 0123456789. $f_\pi$ es una propiedad fugitiva.

Brouwer usa el término $\lambda_f$ para denotar el más pequeño número natural que tenga la propiedad fugitiva $f$, en el caso hipotético de que haya números que la tienen. Digo que $x$ es un infra-número de la propiedad $f$ si $x < \lambda_f$, y que $x$ es un supra-número de $f$ si $\lambda_f \leq x$. (Evidentemente, en el instante mismo en que se demuestre que un cierto número es un supra-número de $f$, $f$ dejará de ser una propiedad fugitiva.) Considérese la secuencia de racionales $a_1$, $a_2$, . . . , definida por:

$$a_k = \left(-\frac{1}{2}\right)^k \quad \text{si } k \text{ es un infra-número de } f$$

$$a_k = \left(-\frac{1}{2}\right)^{\lambda_f} \quad \text{si } k \text{ es un supra-número de } f$$

Es claro que la secuencia converge a un límite $p_f$. También es claro que, si $f$ es una propiedad fugitiva, tenemos que $p_f$ no es igual a 0, pero tampoco es diferente de 0.

Hay importantes teoremas del análisis matemático que no pueden demostrarse sin los recursos rechazados por Brouwer. Doy un ejemplo: *Toda función continua* $\varphi\colon \mathcal{I} \to \mathbb{R}$, *definida en un intervalo cerrado* $\mathcal{I} \subset \mathbb{R}$, *tiene un máximo; en otras palabras, hay un* $\alpha \in \mathcal{I}$ *y un entorno* $\mathcal{U}$ *de* $\alpha$ *tales que* $\varphi(\alpha) \geq \varphi(\xi)$ *para todo* $\xi \in \mathcal{U} \cap \mathcal{I}$. Brouwer (1924; CW I, p. 271) propone el siguiente caso en el cual, según él, este pretendido teorema no se cumple: Enúmerense las fracciones irreducibles, mayores que 0 y menores que 1, cuyo denominador es una potencia de 2, en el orden siguiente: (*a*) la fracción con menor denominador precede a la con denominador mayor; (*b*) cuando los denominadores son iguales, la fracción con menor numerador precede a la con numerador mayor. Las fracciones duales así ordenadas constituyen una secuencia de Cauchy $\delta_1, \delta_2, \ldots$ Sea $\lambda_\pi$ el más pequeño número natural que tenga la propiedad fugitiva $f_\pi$ definida arriba. Sea $\varphi_n\colon [0,1] \to \mathbb{R}$ una función lineal en $[0,\delta_n)$ y en $(\delta_n,1]$, tal que $\varphi_n(0) = \varphi_n(1) = 0$ y $\varphi_n(\delta_n) = 2^{-n}$. Sea $g_n(x) = \varphi_n(x)$ si $n = \lambda_\pi$ y $g_n(x) = 0$ si $n \neq \lambda_\pi$. Entonces la función $g$, definida por $g(x) = \sum_{n=1}^{\infty} g_n(x)$, es una función continua, definida en el intervalo cerrado [0,1], que no tiene un máximo.[55]

55 Largeault 1992 y Mancosu 1998 son libros introductorios que recomiendo calurosamente a quien desee saber más sobre el intuicionismo.

Solución del ejercicio en la página 501.

El alcance de $\neg$ en su primera posición es la fila $\forall x_1(P^3x_1x_2x_3 \vee \exists x_2\neg P^2x_2f^2x_1x_3)$; el de $\forall x_1$ es $(P^3x_1x_2x_3 \vee \exists x_2\neg P^2x_2f^2x_1x_3)$; el de $P^3$ es $x_1x_2x_3$; el de $\exists x_2$ es $\neg P^2x_2f^2x_1x_3$; el de $\neg$ en su segunda posición es $P^2x_2f^2x_1x_3$; el de $f^2$ es $x_1x_3$, y el de $\vee$ es el par de filas $\langle P^3x_1x_2x_3, \exists x_2\neg P^2x_2f^2x_1x_3\rangle$.

# GLOSARIO

El signo † indica que la palabra así señalada es materia de otro artículo de este glosario.

El artículo **Aplicación** explica varios términos utilizados en el libro y en los otros artículos.

**Aplicación** (alemán, *Abbildung*; francés, *application*; inglés, *mapping*). Sean $G$ y $H$ dos conjuntos cualesquiera. Una *aplicación* $f$ de $G$ en $H$ asigna a cada elemento $x \in G$ un único elemento $f(x) \in H$. Decimos que $G$ es el *dominio* y $H$ el *codominio* de la aplicación $f$. $f(x)$ es el *valor* de $f$ en el *argumento* $x$. El *alcance* de $f$ es el conjunto $V \subseteq H$ de los valores de $f$. Si el alcance $V$ coincide con el codominio $H$, decimos que $f$ aplica $G$ *sobre* $H$ y que $f$ es una aplicación *epiyectiva* o una *epiyección* (F. *surjection*). Decimos que $f$ es una aplicación *inyectiva* o una *inyección* (F. *injection*) si tiene valores diferentes en cada par de argumentos diferentes (esto es, si para todo $x, y \in G$, $x \neq y$ implica que $f(x) \neq f(y)$). Si $f$ es a la vez inyectiva y epiyectiva, decimos que es una aplicación *biyectiva* o una *biyección* (F. *bijection*). Designamos la aplicación $f$ de $G$ en $H$ mediante la expresión simbólica $f: G \rightarrow H$. La expresión simbólica $f: x \mapsto f(x)$ dice que $f$ asigna el valor $f(x)$ al argumento $x$. Si $A \subseteq G$, el conjunto $\{f(x): x \in A\}$ se llama la *imagen* de $A$ por $f$ y se designa $f(A)$. Si $B \subseteq H$, el conjunto $\{x: f(x) \in B\}$ se llama la *preimagen* de $B$ por $f$ y se designa $f^{-1}(B)$.

Si $f: G \rightarrow H$ es una aplicación y $F \subseteq G$, la *restricción* de $f$ a $F$ (simbolizada $f|F$) es la aplicación $f_1: F \rightarrow H$, definida por $f_1(x) = f(x)$ para todo $x \in F$. Por otro lado, si $f: G \rightarrow H$ es una aplicación y $G \subseteq K$, cualquier aplicación $f_2: K \rightarrow H$ definida de tal modo que $f = f_2|G$ es una *extensión* de $f$ a $K$.

Considérense dos aplicaciones, $f: G \rightarrow H$ y $g: H \rightarrow K$. Entonces, la aplicación $g \circ f: G \rightarrow K$, $x \mapsto g(f(x))$ es la *aplicación compuesta* de $f$ por $g$, la cual tiene, para cada $x \in G$, un valor idéntico al valor de $g$ en $f(x)$. A veces, en vez de $g \circ f$, escribimos simplemente $gf$.

El *grafo* de la aplicación $f\colon G \to H$; $x \mapsto f(x)$ es el conjunto de pares ordenados $\{\langle x,y\rangle\colon x \in G, y = f(x) \in H\}$. Mediante el concepto de grafo se suele dar la siguiente definición conjuntista de aplicación: Una aplicación $f$: $G \to H$; $x \mapsto f(x)$ es un trío ordenado $\langle Q,G,H\rangle$, donde $Q = \{\langle x,f(x)\rangle\colon x \in G, f(x) \in H\}$ y, para cada $x$, $y \in G$, $f(x) \neq f(y)$ sólo si $x \neq y$.

Una aplicación $f\colon G \to H$ cuyo dominio y codominio son conjuntos de números (naturales, enteros, racionales, reales o complejos) suele llamarse *función*.

Supongamos que hay una relación $n$-aria $R_G$ entre los miembros del conjunto $G$. La aplicación $f$: $G \to H$ *induce* en $H$ la relación $n$-aria $R_{fG}$ definida como sigue: Si $h_1, \ldots, h_n \in H$, $R_{fG}(h_1, \ldots, h_n)$ si y sólo si $h_1 = f(g_1), \ldots, h_n = f(g_n)$ y $R_G(g_1, \ldots, g_n)$. Si $R_{fG}$ coincide con una relación $n$-aria preexistente en $H$ solemos dar a ésta el mismo nombre que a $R_G$ y decir que la aplicación $f$ *preserva* esa relación. Por ejemplo, la función $x \mapsto x^2$ definida en el dominio de los números naturales preserva la relación binaria "mayor que", puesto que $x < y$ si y sólo si $x^2 < y^2$. (No así, en cambio, la función $x \mapsto x^2$ definida en el domino de los enteros, puesto que $-x < -y$ implica que $(-x)^2 > (-y)^2$.)

**Boole**, **álgebra de.** Consíderese un conjunto no vacío $\mathscr{B}$, en que se han distinguido dos objetos $\mathbf{0}$ y $\mathbf{1}$ y se han definido una †operación algebraica unaria $\mathsf{C}$ y dos operaciones binarias $\cap$ y $\cup$. Sean $a$, $b$ y $c$ elementos cualesquiera de $\mathscr{B}$. La estructura $\langle \mathscr{B},\mathbf{0},\mathbf{1},\mathsf{C},\cap,\cup\rangle$ constituye un *álgebra de Boole* si cumple con las siguientes condiciones **B1–B11**. Como es habitual, escribo $a'$ en vez de $\mathsf{C}(a)$, y $a \cap b$, $a \cup b$ en vez de $\cap(a,b)$, $\cup(a,b)$.

**B1** $a \cap b = b \cap a$
**B2** $a \cup b = b \cup a$
**B3** $a \cap (b \cap c) = (a \cap b) \cap c$
**B4** $a \cup (b \cup c) = (a \cup b) \cup c$
**B5** $a \cup (b \cap c) = (a \cup b) \cap (a \cup c)$
**B6** $a \cap (b \cup c) = (a \cap b) \cup (a \cap c)$
**B7** $a \cup (a \cap b) = (a \cup b) \cap a = a$
**B8** $a \cup \mathbf{0} = a$
**B9** $a \cup \mathbf{1} = \mathbf{1}$
**B10** $a \cup a' = \mathbf{1}$
**B11** $a \cap a' = \mathbf{0}$

Las operaciones $\cup$ y $\cap$ son pues conmutativas (**B1**, **B2**) y asociativas (**B3**, **B4**). Además, cada una es distributiva con respecto a la otra (**B5**, **B6**). Las condiciones **B1**–**B7** implican que las operaciones $\cup$ y $\cap$ son idempotentes ($a \cup a = a$ y $a \cap a = a$). Se puede demostrar que el álgebra de Boole $\langle\mathcal{B},\mathbf{0},\mathbf{1},\mathsf{C},\cap,\cup\rangle$ obedece a las *leyes de De Morgan*: $(a \cup b)' = a' \cap b'$ y $(a \cap b)' = a' \cup b'$. Diremos que $a$ es menor o igual que $b$ ($a \leq b$) si y sólo si $a \cup b = b$. $\langle\mathcal{B},\leq\rangle$ constituye un †orden parcial. En virtud de **B8** y **B9**, $\mathbf{0} = \min \mathcal{B}$ y $\mathbf{1} = \max \mathcal{B}$. Como $a \cup (a \cup b) = (a \cup a) \cup b = a \cup b$, es claro que $a \leq a \cup b = b \cup a$. Por otra parte, con arreglo a **B7**, $(a \cap b) \cup a = a$, de modo que $a \cap b \leq a$.

Sea $K$ un conjunto cualquiera, y $\mathcal{P}K$ su conjunto potencia (p. 3). Designamos con $A \cap B$ la *intersección* de $A$ y $B$, esto es, el conjunto $\{x{:}x \in A$ y $x \in B\}$; con $A \cup B$ la *unión* de $A$ y $B$, esto es, el conjunto $\{x{:}x \in A$ o $x \in B\}$; y con $\mathsf{C}_K(A)$ el *complemento* de $A$ en $K$, esto es, el conjunto $K\setminus A = \{x{:}x \in K$ y $x \notin A\}$. La estructura $\langle\mathcal{P}K,\varnothing,K,\mathsf{C}_K,\cap,\cup\rangle$ es un álgebra de Boole. (Demuéstrelo el lector). Comúnmente, el contexto fija el conjunto $K$ con respecto al cual se toma el complemento, y escribimos $A'$ en vez de $\mathsf{C}_K(A)$;

**Cuerpo** (alemán, *Körper*; francés, *corps*; inglés, *field*)**.** Sea $\langle\mathcal{K},\oplus\rangle$ un †grupo abeliano, con elemento neutro $\mathbf{0} \in \mathcal{K}$. Suponemos que $\mathcal{K}$ contiene por lo menos un elemento distinto de $\mathbf{0}$. Sea $\otimes{:}\mathcal{K} \times \mathcal{K} \rightarrow \mathcal{K}$ una †aplicación tal que (i) $\langle\mathcal{K}\setminus\{\mathbf{0}\},\otimes\rangle$ es un grupo abeliano con elemento neutro $\mathbf{1}$,[1] (ii) para cualquier $k \in \mathcal{K}$, $k \otimes \mathbf{0} = \mathbf{0} \otimes k = \mathbf{0}$, y (iii) cualesquiera que sean $a,b,c \in \mathcal{K}$, $a \otimes (b \oplus c) = (a \otimes b) \oplus (a \otimes c) = (b \oplus c) \otimes a$. Entonces, $\langle\mathcal{K},\mathbf{0},\mathbf{1},\oplus,\otimes\rangle$ es un *cuerpo*. Sea $a \in \mathcal{K}$. Si $a \neq \mathbf{0}$, $a$ tiene dos inversos: uno por $\oplus$, que denotamos con $-a$, y uno por $\otimes$ que denotamos con $a^{-1}$. Si $a = \mathbf{0}$, obviamente, es su propio inverso por $\oplus$ y no tiene un inverso por $\otimes$.

El lector comprobará fácilmente que, si $\mathcal{K}$ es el conjunto de todas las fracciones (propias e impropias), $\mathbf{0}$ y $\mathbf{1}$ son el cero y el uno, y $\oplus$ y $\otimes$ son, respectivamente, la adición y la multiplicación de fracciones, $\langle\mathcal{K},\mathbf{0},\mathbf{1},\oplus,\otimes\rangle$ es un cuerpo: el *cuerpo de los racionales*, habitualmente llamado $\mathbb{Q}$.

[1] $\mathcal{K}\setminus\mathcal{B}$ designa el conjunto de todos los elementos de $\mathcal{K}$ que no pertenecen a $\mathcal{B}$. Por lo tanto, $\mathcal{K}\setminus\{\mathbf{0}\}$ es el conjunto de todos los elementos de $\mathcal{K}$ con excepción de $\mathbf{0}$. En rigor, habría que formular la condición (i) así: Sea $\otimes'$ la restricción de $\otimes$ a $(\mathcal{K}\setminus\{\mathbf{0}\})^2$; $\langle\mathcal{K}\setminus\{\mathbf{0}\},\otimes'\rangle$ es un grupo abeliano con elemento neutro $\mathbf{1}$. Entonces, la regla (ii) caracteriza a $\otimes$ como una *extensión* de $\otimes'$ a $\mathcal{K}^2$.

El cuerpo $\mathbb{R}$ de los reales se suele definir como una extensión de $\mathbb{Q}$, pero aquí daré la definición, menos intuitiva pero mucho más elegante, propuesta por David Hilbert en 1900. Sea $\mathbb{K} = \langle \mathcal{K}, \mathbf{0}, \mathbf{1}, \oplus, \otimes \rangle$ un cuerpo cualquiera. Supongamos que $\mathcal{K}$ incluye una parte no vacía $P \subseteq \mathcal{K}$ tal que (i) si $a \in P$ y $b \in P$, $a \oplus b \in P$ y $a \otimes b \in P$, y (ii) si $a \in \mathcal{K}$, $a$ cumple con una y sólo una de las tres condiciones siguientes: $a \in P$, $a = \mathbf{0}$, o $-a \in P$. En tal caso, decimos que $P$ es el conjunto de los elementos *positivos* de $\mathbb{K}$ y que $\mathbb{K}$ es un cuerpo *ordenado*. Esta denominación se justifica porque la existencia de $P$ determina en $\mathbb{K}$ la relación de †orden lineal < definida por: $a < b$ (léase: "$a$ es menor que $b$") si y sólo si $b \oplus -a \in P$. Como es habitual, escribimos '$a \leq b$' por '$a < b$ o $a = b$'. Decimos que un elemento $a \in \mathcal{K}$ es una *cota superior* del conjunto $C \subseteq \mathcal{K}$ si $c \leq a$ para todo $c \in C$. Si existe una cota superior de $C$, decimos que $C$ es un conjunto *acotado por arriba*. En particular, decimos que $a_0$ es el *supremo* o *cota superior mínima* de $C$ si $a_0$ es una cota superior de $C$ y cualquier *otra* cota superior de $C$ es mayor que $a_0$. (Los conceptos de conjunto *acotado por abajo*, *cota inferior* e *ínfimo* o *cota inferior máxima* se definen en forma análoga). Decimos que el cuerpo ordenado $\mathbb{K}$ es *completo* si todo conjunto $C \subseteq \mathcal{K}$ acotado por arriba tiene una cota superior mínima. Decimos que el cuerpo ordenado $\mathbb{K}$ es *arquimédico* si, cualesquiera que sean los elementos $a, b \in \mathbb{K}$, si $a \in P$ (esto es, si $\mathbf{0} < a$) siempre existe un número natural $n$ tal que $b$ es menor que $n$ veces $a$ (vale decir, $b < a \oplus a \oplus \ldots \oplus \mathbf{0}$, con $\oplus$ repetido $n$ veces). Se puede demostrar que si $\mathbb{K}$ y $\mathbb{K}'$ son dos cuerpos ordenados, completos y arquimédicos hay un †isomorfismo $f: \mathbb{K} \to \mathbb{K}'$, esto es, una biyección que preserva todas las propiedades estructurales de $\mathbb{K}$ (por ejemplo: $f(a \otimes b) = f(a) \otimes f(b)$, $f(P) = P'$, etc.). Por lo tanto, es lícito hablar de una estructura única, $\mathbb{R}$, realizada por todo *cuerpo ordenado, completo y arquimédico*. $\mathbb{R}$ se llama *el cuerpo de los reales*. (Obsérvese que $\mathbb{Q}$ es un cuerpo ordenado arquimédico, pero no completo: el conjunto de todas las fracciones cuyo cuadrado es mayor que 2 está acotado por abajo, pero no tiene una cota inferior máxima).

Dado el cuerpo $\mathbb{R}$ de los reales, el cuerpo $\mathbb{C}$ de los complejos puede definirse como sigue. Un elemento de $\mathbb{C}$ —un *número complejo*— es cualquier par ordenado de números reales: $\langle a,b \rangle \in \mathbb{C}$ si y sólo si $a, b \in \mathbb{R}$. En vez de $\langle a,b \rangle$ escribimos $(a,b)$. La adición y la multiplicación de complejos se define como sigue: $(a,b) + (c,d) = (a+b,c+d)$, $(a,b) \times (c,d) = (ac+bd,ad-bc)$, donde los signos +, – y × que figuran *dentro* de los parénte-

sis designan, respectivamente, la adición, la subtracción (esto es, la adición del inverso) y la multiplicación en $\mathbb{R}$. En $\mathbb{C}$, el elemento neutro de la adición es (0,0) y el elemento neutro de la multiplicación es (1,1). El lector puede tratar de probar que la estructura así definida es un cuerpo.

**Equivalencia**. Sea $\mathscr{C}$ un conjunto cualquiera. Considérese una relación $R$ entre pares de elementos de $\mathscr{C}$ tal que (i) $R$ es *reflexiva*: cada elemento de $\mathscr{C}$ la tiene consigo mismo (simbólicamente: $xRx$); (ii) $R$ es *simétrica*: si $a$ tiene la relación $R$ con $b$, $b$ tiene $R$ con $a$ ($aRb \Rightarrow bRa$); y (iii) $R$ es *transitiva*: si $aRb$ y $bRc$, entonces $aRc$. Si la relación $R$ cumple estas condiciones decimos que $R$ es una *equivalencia*. Dada una equivalencia $R$ definida en un conjunto $\mathscr{C}$, los elementos de $\mathscr{C}$ pueden clasificarse en *clases de equivalencia*: $a$ y $b$ pertenecen a la misma clase de equivalencia si y sólo si $aRb$. Esta clasificación es exhaustiva, puesto que, en virtud de (i) todo elemento de $\mathscr{C}$ tiene $R$ con algo; además, en virtud de (iii), las clases son mutuamente exclusivas. Por lo tanto, la clasificación de $\mathscr{C}$ en clases de equivalencia constituye una †*partición*.

**Espacio topológico.** Sea $\mathscr{E}$ un conjunto cualquiera. Decimos que $\mathscr{T} \subseteq \mathscr{P}(\mathscr{E})$ es una *topología* en $\mathscr{E}$ si cumple las cuatro condiciones siguientes:

(i) $\mathscr{E} \in \mathscr{T}$;
(ii) $\varnothing \in \mathscr{T}$;
(iii) si $X$ e $Y$ pertenecen a $\mathscr{T}$, la intersección $X \cap Y$ también pertenece a $\mathscr{T}$;
(iv) si $X_1, X_2, \ldots$ es una lista (posiblemente infinita) de elementos de $\mathscr{T}$, la unión $\bigcup_{k\in\mathbb{N}} X_k$ de todos los elementos de la lista también es un elemento de $\mathscr{T}$.

Un *espacio topológico* es un par $\langle\mathscr{E},\mathscr{T}\rangle$ tal que $\mathscr{E}$ es un conjunto no vacío y $\mathscr{T}$ es una topología en $\mathscr{E}$. Llamamos *puntos* a los elementos de $\mathscr{E}$ y *abiertos* a los elementos de $\mathscr{T}$. Si $X$ es un abierto, decimos que su complemento $\mathscr{E}\setminus X$ es un *cerrado*. Este vocabulario se comprenderá mejor a la luz de las siguientes explicaciones. Si $\langle\mathscr{E},\mathscr{T}\rangle$ es un espacio topológico y $x$ es un punto de $\mathscr{E}$, cualquier subconjunto de $\mathscr{E}$ que contenga a $x$ se llama un *entorno* de $x$. Sea $x \in \mathscr{E}$ y $X \subseteq \mathscr{E}$. Entonces ($a$) $x$ es un *punto de acumulación* (o *punto*

*límite*) de $X$ si todo entorno de $x$ contiene a lo menos un punto de $X$; (*b*) $x$ es un *punto fronterizo* de $X$ si todo entorno de $x$ contiene a lo menos un punto de $X$ y un punto de $\mathcal{E}\setminus X$; (*b*) $x$ es un *punto interior* de $X$ si por lo menos un entorno abierto de $x$ (esto es, un subconjunto de $\mathcal{E}$ que contiene a $x$ y pertenece a $\mathcal{T}$) no contiene ningún punto que no pertenezca a $X$ (según esta definición, $x$ sólo puede ser un punto interior de $X$, si $x \in X$). La unión de $X$ y el conjunto todos sus puntos de acumulación se llama la *clausura* de $X$. El conjunto de todos los puntos fronterizos de $X$ se llama la *frontera* de $X$. El conjunto de todos los puntos interiores de $X$ se llama el *interior* de $X$. Se puede demostrar fácilmente que todo abierto es idéntico a su propio interior y que todo cerrado es idéntico a su propia clausura; además, la clausura de cualquier conjunto $X \subseteq \mathcal{E}$ es igual a la unión de su interior y su frontera.

Sean $\langle\mathcal{E},\mathcal{T}\rangle$ y $\langle\mathcal{E}',\mathcal{T}'\rangle$ dos espacios topológicos y $f$ una †aplicación de $\mathcal{E}$ en $\mathcal{E}'$. Decimos que $f$ es una aplicación *continua* si y sólo si la pre-imagen por $f$ de cada abierto de $\langle\mathcal{E}',\mathcal{T}'\rangle$ es un abierto de $\langle\mathcal{E},\mathcal{T}\rangle$. En otras palabras, $f{:}\mathcal{E} \to \mathcal{E}'$ es continua si y sólo si, para cada $X \in \mathcal{T}'$, $f^{-1}(X) \in \mathcal{T}$. Si la aplicación continua $f{:}\mathcal{E} \to \mathcal{E}'$ es biyectiva y su inversa $f^{-1}{:}\mathcal{E}' \to \mathcal{E}$ también es continua, $f$ es un *homeomorfismo*, esto es, un †isomorfismo de espacios topológicos, y $\langle\mathcal{E},\mathcal{T}\rangle$ y $\langle\mathcal{E}',\mathcal{T}'\rangle$ son espacios *homeomórficos* (vale decir, indistinguibles en cuanto espacios topológicos).

**Grupo.** Sea $\mathcal{G}$ un conjunto no vacío, $e$ un elemento determinado de $\mathcal{G}$, $\mathbf{i}$ una †operación algebraica unaria en $\mathcal{G}$ y $\otimes$ una operación algebraica binaria en $\mathcal{G}$. $\langle\mathcal{G},e,\mathbf{i},\otimes\rangle$ es un *grupo* si se cumplen las tres condiciones siguientes: (i) $\otimes$ es asociativa: si $x, y, z \in \mathcal{G}$, $x \otimes (y \otimes z) = (x \otimes y) \otimes z$; (ii) para cada $x \in \mathcal{G}$, $x \otimes e = e \otimes x = x$; (iii) para cada $x \in \mathcal{G}$, $x \otimes \mathbf{i}(x) = \mathbf{i}(x) \otimes x = e$. La operación $\otimes$ se llama el *producto* del grupo, $e$ es el *elemento neutro*, $\mathbf{i}(x)$ es el *elemento inverso* de $x$. Normalmente, se escribe $xy$ en vez de $x \otimes y$ y el elemento inverso de $x$ se llama $x^{-1}$ en vez de $\mathbf{i}(x)$; en el caso de ciertos grupos es tradición escribir $x + y$ por $x \otimes y$ y llamar $-x$ al inverso de $x$.

**Grupo abeliano.** El grupo $\langle\mathcal{G},e,\mathbf{i},\otimes\rangle$ es *abeliano* si la operación $\otimes$ es conmutativa, esto es, si $(x \otimes y) = (y \otimes x)$ para cada par de elementos si $x, y \in \mathcal{G}$.

**Isomórfico, isomorfismo.** Sea $f$ una aplicación biyectiva de un conjunto $\mathcal{A}$ en un conjunto $\mathcal{B}$. Si $\mathcal{A}$ tiene una estructura —digamos, un orden parcial, o

una estructura de cuerpo o de espacio topológico— $f$ induce en $\mathcal{B}$ una estructura de la misma clase. Por ejemplo, si $\mathcal{A}$ contiene en virtud de su estructura un elemento distinguido $e$, $f(e)$ es un elemento distinguido de $\mathcal{B}$ en virtud de la estructura inducida por $f$; si la estructura de $\mathcal{A}$ está definida (en parte) por una relación diádica $R$ y $aRb$, la estructura inducida estará definida (en parte) por una relación diádica $R^*$ tal que $f(a)R^*f(b)$; si la estructura de $\mathcal{A}$ incluye una operación binaria que asigna el objeto $z \in \mathcal{A}$ al par de objetos $x, y \in \mathcal{A}$, la estructura inducida asigna precisamente el objeto $f(z)$ al par de objetos $f(x)$ y $f(y)$; si la estructura de $\mathcal{A}$ selecciona una clase $\mathcal{T}$ de subconjuntos de $\mathcal{A}$, la estructura inducida selecciona una clase $\mathcal{T}^*$ de subconjuntos de $\mathcal{B}$ tal que $X \in \mathcal{T}$ si y sólo si $f(X) \in \mathcal{T}^*$. Supongamos ahora que tanto $\mathcal{A}$ como $\mathcal{B}$ tienen una estructura y que la estructura de $\mathcal{B}$ coincide con la estructura inducida por $f$. Decimos entonces que la biyección $f$ es un *isomorfismo*. En tal caso, obviamente, la biyección inversa $f^{-1}$ también es un isomorfismo. Dos conjuntos estructurados $\mathcal{A}$ y $\mathcal{B}$ son *isomórficos* si existe un isomorfismo entre ellos.

**Operación algebraica.** Si $\mathcal{C}$ es un conjunto cualquiera, llamamos $\mathcal{C}^n$ al conjunto de todos los $n$-tuplos ordenados que pueden formarse con elementos de $\mathcal{C}$ ($\mathcal{C}^n = \{\langle x_1, \ldots, x_n\rangle : x_k \in \mathcal{C}, 1 \leq k \leq n\}$. Una *operación algebraica n-aria* en un conjunto $\mathcal{C}$ es sencillamente una aplicación de $\mathcal{C}^n$ en $\mathcal{C}$. En vez de 'operación algebraica', solemos decir 'operación'. Si $f{:}\mathcal{C}^2 \to \mathcal{C}$ es una operación binaria matemáticamente significativa (como la suma o la multiplicación de enteros, etc.) es corriente simbolizarla mediante un ideograma *ad hoc* (+, ×, etc.) que se *intercala* entre los miembros del par ordenado a que se aplica la operación: escribimos '5 + 3', en vez de $+(\langle 5,3\rangle)$. Consideremos ahora una operación binaria $*$ en un conjunto $\mathcal{C}$. Decimos que $*$ es *conmutativa* si para cualesquiera objetos $x, y \in \mathcal{C}$, $x * y = y * x$. Decimos que $*$ es *asociativa* si, para cualesquiera objetos $x, y, z \in \mathcal{C}$, $x * (y * z) = (x * y) * z$. Si $\bullet$ es otra operación binaria en $\mathcal{C}$ decimos que $*$ es *distributiva con respecto a* $\bullet$ si, para cualesquiera objetos $x, y, z \in \mathcal{C}$, $x * (y \bullet z) = (x * y) \bullet (x * z)$.

**Orden.** Considérese un conjunto $\mathcal{C}$ y una relación diádica $\leq$ entre pares ordenados de elementos de $\mathcal{C}$. Para indicar que $a \in \mathcal{C}$ tiene la relación $\leq$ con

$b \in \mathscr{C}$ escribimos '$a \leq b$' (en vez de $\leq(a,b)$). Diremos que $\leq$ define un *orden parcial* en $\mathscr{C}$ y que $\mathscr{C}$ está *parcialmente ordenado* por $\leq$, si se cumplen las condiciones siguientes:

(i) la relación $\leq$ es *reflexiva*: si $a \in \mathscr{C}$, $a \leq a$;
(ii) la relación $\leq$ es *antisimétrica*: $a \leq b$ y $b \leq a$ sólo si $a = b$;
(iii) la relación $\leq$ es *transitiva*: si $a \leq b$ y $b \leq c$, entonces $a \leq c$.

Decimos que $\leq$ define un *orden total* en $\mathscr{C}$ y que $\mathscr{C}$ está *totalmente ordenado* por $\leq$, si además de las condiciones (i), (ii) y (iii) se cumple la condición siguiente:

(iv) si $a, b \in \mathscr{C}$, entonces siempre ocurrirá que $a \leq b$ o que $b \leq a$.

En vez de 'orden total' suele decirse 'orden simple' u 'orden lineal'. Dado un orden total definido por la relación $\leq$, la relación diádica $<$ está definida por la condición: $a < b$ si y sólo si $a \leq b$ y $a \neq b$. La relación $<$ es *asimétrica*: $a < b$ implica que es falso que $b < a$. El concepto de orden total puede definirse también en términos de una relación transitiva y asimétrica $<$ que satisfaga el requisito de *tricotomía*: si $a, b \in \mathscr{C}$, siempre ocurrirá que $a < b$, o que $b < a$, o que $a = b$. En tal caso, la relación $\leq$ definida por la condición '$a \leq b$ si y sólo si $a < b$ o $a = b$' define en $\mathscr{C}$ un orden total de acuerdo con nuestra primera definición.

Sea $\mathscr{C}$ un conjunto parcialmente ordenado por la relación $\leq$. Sea $U \subseteq \mathscr{C}$. Decimos que $u$ es un elemento *maximal* de $U$ si $u \in U$ y no existe un $v \in U$ tal que $u < v$. Decimos que $u$ es un elemento *minimal* de $U$ si $u \in U$ y no existe un $v \in U$ tal que $v < u$. Un elemento $a \in \mathscr{C}$ es una *cota superior* de $U$ si todo $x \in U$ cumple la condición $x \leq a$. $a$ es una *cota inferior* de $U$ si todo $x \in U$ cumple la condición $a \leq x$. Si $U$ tiene una cota superior, decimos que está *acotado por arriba*; si tiene una cota inferior, decimos que está *acotado por abajo*; si tiene una cota superior y una cota inferior, decimos simplemente que $U$ es un conjunto acotado. Obsérvese que a lo sumo una cota superior de $U$ puede pertenecer a $U$. Si tal cota superior existe la llamamos el *máximo* de $U$ (abreviado: max $U$). Obviamente, si max $U$ existe, es un elemento maximal de $U$ (en efecto, el único). Por otra parte, no más de una cota inferior de $U$ puede pertenecer a $U$. Si tal cota inferior existe, la llamamos el *mínimo* de $U$ (abreviado: min $U$). Si min $U$ existe, es un

elemento minimal de $U$ (en efecto, el único). Supongamos ahora que $U$ tiene una cota superior $p$ tal que, si $q$ es una cota superior de $U$, $p \leq q$. Obviamente, no puede haber más de un objeto $p$ con esta propiedad. Si tal objeto existe, lo llamamos la *cota superior mínima* o el *supremo* de $U$ (abreviado: sup $U$). Del mismo modo, la *cota inferior máxima* o *ínfimo* de $U$ (abreviado: inf $U$) es el objeto $r$ tal que $r$ es una cota inferior de $U$ y si $r$ es cualquier cota inferior de $U$, $s \leq r$; obviamente, si inf $U$ existe, es único.

**Partición**. Sea $\mathcal{C}$ un conjunto y $\mathcal{P}(\mathcal{C})$ el conjunto de los subconjuntos (las partes) de $\mathcal{C}$. Decimos que $\mathsf{P} \subseteq \mathcal{P}(\mathcal{C})$ constituye una *partición* de $\mathcal{C}$ si cada elemento de $\mathcal{C}$ pertenece a *uno y sólo uno* de los conjuntos contenidos en $\mathsf{P}$. Si $\mathsf{P}$ es una partición de $\mathcal{C}$ y $x \in \mathcal{C}$, el elemento de $\mathsf{P}$ al que pertenece $x$ suele llamarse $[x]$. Sea $R$ la relación entre pares de elementos de $\mathcal{C}$ definida por la condición: $xRy$ si y sólo si $x \in [y]$. En tal caso, $[y]$ es el único elemento de $\mathsf{P}$ a que pertenece $x$, esto es, $[y] = [x]$. Por ende, si $x \in [y]$ y $z \in [x]$, $z \in [y]$; esto es, si $zRx$ y $xRy$, entonces $zRy$: $R$ es una relación transitiva. Como obviamente $y \in [y]$, tenemos además que si $xRy$, $yRx$: $R$ es reflexiva y simétrica. En otras palabras, $R$ es una †equivalencia. Hemos demostrado así que toda partición de un conjunto $\mathcal{C}$ define una equivalencia entre los elementos de $\mathcal{C}$.

**Producto cartesiano.** Sean $A$ y $B$ dos conjuntos cualesquiera. El *producto cartesiano* $A \times B$ es el conjunto de todos los pares ordenados $\langle a,b \rangle$ tales que $a \in A$ y $b \in B$. Si $C$ es un tercer conjunto, $A \times B \times C = (A \times B) \times C$. El producto cartesiano de $n$ conjuntos se define en forma análoga. Como es obvio, también se puede formar el producto cartesiano de un conjunto consigo mismo: $A \times A = A^2 = \{\langle x,y \rangle : x, y \in A\}$. Repitiendo esta operación dos, tres, . . . , $n$–1 veces, se obtienen los productos cartesianos $A^3$, $A^4$, $A^n$. $A^n$ es llamado a veces el $n$-ésimo producto cartesiano de $A$ (entiéndase: consigo mismo).

# OBRAS CITADAS

Ackermann, W. (1924). "Begründung des Tertium non datur mittels der Hilbertschen Theorie der Widerspruchsfreiheit". *Mathematische Annalen*. **92**: 1–35.

Ackermann, W. (1928). "Zum Hilbertschen Aufbau der reellen Zahlen". *Mathematische Annalen*. **99**: 118–133.

Ackermann, W. (1940). "Zur Widerspruchsfreiheit der Zahlentheorie". *Mathematische Annalen*. **117**: 162–194.

Ash, J. M. (1989). "Uniqueness of representation by trigonometric series". *American Mathematical Monthly*. **96**: 873–885.

Baker, G. P. y P. M. S. Hacker (1984). *Frege: Logical Excavations*. Oxford: Blackwell.

Baldwin, J. M., ed. (1901-1905). *Dictionary of Philosophy and Psychology*. New York: Macmillan. 3 vols. in 4.

Becker, O. (1954). *Grundlagen der Mathematik in ihrer geschichtlichen Entwicklung*. Freiburg i. Br.: Karl Alber.

Behmann, H. (1922). "Beiträge zur Algebra der Logik, insbesondere zum Entscheidungsproblem". *Mathematische Annalen*. **86**: 163–229.

Benacerraf, P. y H. Putnam, eds. (1983). *Philosophy of Mathematics: Selected Readings*. Second Edition. Cambridge: Cambridge University Press.

Bernays, P. (1926). "Axiomatische Untersuchungen des Aussagenkalküls der *Principia Mathematica*". *Mathematische Zeitschrift*. **25**: 305–320.

Bernays, P. (1935). "Hilberts Untersuchungen über die Grundlagen der Arithmetik". En Hilbert GA, vol. III, pp. 196–216.

Bernays, P. y M. Schönfinkel (1928). "Zum Entscheidungsproblem der mathematischen Logik". *Mathematische Annalen*. **99**: 342–372.

Birkhoff, G. y J. v. Neumann (1936). "The logic of quantum mechanics". *Annals of Mathematics*. **37**: 823–843.

Bishop, E. (1967). *Foundations of Constructive Analysis*. New York: McGraw-Hill.

Bolzano, B. (1964). *Paradoxien des Unendlichen*. Hamburg: Felix Meiner. (primera edición: 1851).

Boole, G. (1847). *The Mathematical Analysis of Logic*. Cambridge: Macmillan.

Boole, G. (1854). *An Investigation of the Laws of Thought on which are Founded the Mathematical Theories of Logic and Probabilities*. London: Macmillan.

Boolos, G. (1987). "The consistency of Frege's *Foundations of Arithmetic*". En Demopoulos 1995, pp. 211-233. (Originalmente publicado en J. J. Thomson, ed., *On Being and Saying: Essays for Richard Cartwright*, Cambridge MA, MIT Press, 1987, pp. 3-20).

Boolos, G. (1990). "The standard of equality of numbers". En Demopoulos 1995, pp. 234-254. (Originalmente publicado en *Meaning and Method: essays in honor of Hilary Putnam*, Cambridge, Cambridge University Press, 1990, pp. 261-277).

Boolos, G. y R. Jeffrey (1974). *Computability and Logic*. Cambridge: Cambridge University Press.

Boolos, G. y R. Jeffrey (1989). *Computability and Logic*. Third Edition. Cambridge: Cambridge University Press.

Bourbaki, N. (1970). *Théorie des ensembles*. Paris: Hermann.

Browder, F. E., ed. (1976). *Mathematical Developments Arising from Hilbert Problems*. Proceedings of the Symposium in Pure Mathematics of the American Mathematical Society held at Northern Illinois University, Dekalb, Illinois in May 1974. Providence, RI: American Mathematical Society. Proceedings of Symposia in Pure Mathematics, XXVIII.

Brouwer, L. E. J. (CW). *Collected Works*. Volume 1, Philosophy and the Foundations of Mathematics. Edited by A. Heyting. Amsterdam: North-Holland, 1975.

Brouwer, L. E. J. (1905). *Leven, Kunst en Mystiek*. Delft: Waltman. (Extractos, traducidos al inglés, en Brouwer, CW I, 1-10; traducción inglesa completa por W. P. van Stigt en *Notre Dame Journal of Formal Logic*, **37** (1996)).

Brouwer, L. E. J. (1908a). "Die möglichen Mächtigkeiten". *Atti IV Congresso Internazionale di Matematici. Bologna*. Pp. 569-571. (Reproducido en Brouwer, CW I, 102-104).

Brouwer, L. E. J. (1908b). "The unreliability of the logical principles". En Brouwer CW, vol. I, pp. 107-111. (Traducción inglesa de Brouwer, "De onbetrouwbaarheid der logische principes", *Tijdschrift voor wijsbegeerte*, **2**: 152-158 (1908)).

Brouwer, L. E. J. (1911). "Beweis des Invarianz des *n*-dimensionalen Gebiets". *Mathematische Annalen*. **71**: 314–319.

Brouwer, L. E. J. (1913). "Über den natürlichen Dimensionsbegriff". *Journal für die reine und angewandte Mathematik*. **142**: 146–152.

Brouwer, L. E. J. (1913a). "Intuitionism and formalism". *Bulletin of the American Mathematical Society*. **20**: 81-96.

Brouwer, L. E. J. (1918). “Begründung der Mengenlehre unabhängig vom logischen Satz vom ausgeschlossenen Dritten. Erster Teil: Allgemeine Mengenlehre”. *K. Nederlandse Akademie van Wetenschappen. Verhandelingen.* 1e Sectie 12, No 5 (1918). (Reproducido en Brouwer, CW, I, 150-190).

Brouwer, L. E. J. (1919). “Begründung der Mengenlehre unabhängig vom logischen Satz vom ausgeschlossenen Dritten. Zweiter Teil: Theorie der Punktmengen”. *K. Nederlandse Akademie van Wetenschappen. Verhandelingen.* 1e Sectie 12, No 7 (1919). (Reproducido en Brouwer, CW, I, 191-221).

Brouwer, L. E. J. (1923). “Begründung der Funktionenlehre unabhängig vom logischen Satz vom ausgeschlossenen Dritten. Erster Teil: Stetigkeit, Messbarkeit, Derivierbarkeit”. *K. Nederlandse Akademie van Wetenschappen. Verhandelingen.* 1e Sectie 13, No 2 (1923). (Reproducido en Brouwer, CW, I, 246-267).

Brouwer, L. E. J. (1924). “Über die Bedeutung des Satzes vom ausgeschlossenen Dritten in der Mathematik, insbesondere in der Funktionentheorie”. *Journal für die reine und angewandte Mathematik*. **154**: 1-7. (Reproducido en Brouwer, CW I, 508-515; publicado originalmente en holandés en 1923; traducción inglesa, “On the significance of the principle of the excluded middle in mathematics, especially in function theory”, en van Heijenoort 1967, pp. 334-341).

Brouwer, L. E. J. (1929). “Mathematik, Wissenschaft und Sprache”. *Monatshefte für Mathematik und Physik*. **36**: 153-164. (Reproducido en Brouwer, CW I, 417-428).

Brouwer, L. E. J. (1948). “Consciousness, philosophy and mathematics”. *Proceedings of the 10th International Congress of Philosophy. Amsterdam*. III, pp. 1235-1249. (Reproducido en Brouwer, CW I, 480-494).

Brouwer, L. E. J. (1952). “Historical background, principles and methods of intuitionism”. *South African Journal of Science*. **49**: 139-146. (Reproducido en Brouwer, CW I, 508-515).

Büchi, J. R. (1962). “Turing machines and the *Entscheidungsproblem*”. *Mathematische Annalen*. **148**: 201-213.

Burali-Forti, C. (1897). “Una questione sui numeri transfiniti”. *Rendiconti del Circolo matematico di Palermo*. **11**: 154–164.

Burali-Forti, C. (1897a). “Sulle classi ben ordinate”. *Rendiconti del Circolo matematico di Palermo*. **11**: 260.

Cantor, G. (GA). *Gesammelte Abhandlungen mathematischen und philosophischen Inhalts*. Nebst einem Lebenslauf Cantors von A. Fraenkel. Mit erläuternden Anmerkungen sowie mit Ergänzungen aus dem Briefwechsel Cantor-Dedekind herausgegeben von Ernst Zermelo. Hildesheim: Olms, 1962. (Reprografischer Nachdruck der Ausgabe Berlin 1932).

Cantor, G. (1870). “Beweis, daß eine für jeden reellen Wert von $x$ durch eine trigonometrische Reihe gegebene Funktion $f(x)$ sich nur auf eine einzige Weise in dieser Form darstellen läßt”. *Journal für die reine und angewandte Mathematik*. **72**: 139–142. (Reproducido en Cantor GA, pp. 80-83.

Cantor, G. (1871). “Notiz zu dem Aufsatze: Beweis, daß eine für jeden reellen Wert von $x$ durch eine trigonometrische Reihe gegebene Funktion $f(x)$ sich nur auf eine einzige Weise in dieser Form darstellen läßt”. *Journal für die reine und angewandte Mathematik*. **73**: 294–296. (Reproducido en Cantor GA, pp. 84-86).

Cantor, G. (1872). “Über die Ausdehnung eines Satzes aus der Theorie der trigonometrischen Reihen”. *Mathematische Annalen*. **5**: 123–132. (Reproducido en Cantor GA, pp. 92-102).

Cantor, G. (1874). “Über eine Eigenschaft des Inbegriffes aller reellen algebraischen Zahlen”. *Journal für die reine und angewandte Mathematik*. **77**: 258–262. (Reproducido en Cantor GA, pp. 115-118).

Cantor, G. (1878). “Ein Beitrag zur Mannigfaltigkeitslehre”. *Journal für die reine und angewandte Mathematik*. **84**: 242–258. (Reproducido en Cantor GA, pp. 119-133).

Cantor, G. (1883). “Über unendliche, lineare Punktmannichfaltigkeiten, Nr. 5”. *Mathematische Annalen*. **21**: 545–591. (Reproducido en Cantor GA, pp. 165-209).

Cantor, G. (1886). “Über die verschiedenen Standpunkte in Bezug auf das aktuelle Unendliche”. *Zeitschrift für Philosophie und philosophische Kritik*. **88**: 224–233. (Reproducido en Cantor GA, pp. 370-377).

Cantor, G. (1887/88). “Mitteilungen zur Lehre vom Transfiniten”. *Zeitschrift für Philosophie und philosophische Kritik*. **91**: 81–125; **92**: 240–265. (Reproducido en Cantor GA, pp. 378-439).

Cantor, G. (1890/91). “Über eine elementare Frage der Mannigfaltigkeitslehre”. *Jahresbericht der Deutschen Mathematiker-Vereinigung*. **1**: 75–78. (Reproducido en Cantor GA, pp. 278-281).

Cantor, G. (1895/97). “Beiträge zur Begründung der transfiniten Mengenlehre”. *Mathematische Annalen*. **45**: 581–512; **49**: 207–246. ((Reproducido en Cantor GA, pp. 282-351).

Cantor, G. (1899). “Carta a Richard Dedekind del 3 de agosto de 1899”. En Cantor GA, pp. 443–447. (Nótese que esta carta se imprimió fundida con otra, del 28 de julio de 1899, a la cual pertenecen sólo los cuatro primeros párrafos de la p. 443; cf. Grattan-Guinness 1974).

Cartan, H. (1937). “Théorie des filtres”. *Comptes rendus de l'Académie des Sciences*. **205**: 595-598.

Chang, C. C. y H. J. Keisler (1990). *Model Theory*. Third edition. Amsterdam: North-Holland..

Church, A. (1936). "An unsolvable problem of elementary number theory". *American Journal of Mathematics*. **58**: 345-363 (1936). (Reproducido en Davis 1965, pp. 89-107. Este artículo fue sometido el 19 de abril de 1935. Se publicó un resumen en el *Bulletin of the American Mathematical Society*, **41**: 333 (1935)).

Church, A. (1936a). "A note on the Entscheidungsproblem". *Journal of Symbolic Logic*. **1**: 40-41, 101-102 (1936). (Reproducido en Davis 1965, pp. 110-115; mis citas remiten a este libro, que funde las dos publicaciones citadas, integrando en la primera la corrección aportada por la segunda).

Church, A. (1976). "Comparison of Russell's resolution of the semantical antinomies with that of Tarski". *Journal of Symbolic Logic*. **41**: 747–760. (Versión revisada en R. L. Martin 1984, pp. 289–306).

Church, A. (1984). "Comparison of Russell's resolution of the semantical antinomies with that of Tarski". En R. L. Martin 1984, pp. 289–306. (Versión revisada de Church 1976).

Chwistek, L. (1922). "Über die Antinomien der Principien der Mathematik". *Mathematische Zeitschrift*. **14**: 236–243.

Cocchiarella, N. (1980). "The development of the theory of logical types and the notion of a logical subject in Russell's early philosophy". *Synthese*. **45**: 71–115.

Coffa, J. A. (1981). "Russell and Kant". *Synthese*. **46**: 247-263.

Coffa, J. A. (1991). *The Semantic Tradition from Kant to Carnap: To the Vienna Station*. Edited by L. Wessels. Cambridge: Cambridge University Press.

Copi, I. M. (1950). "The inconsistency or redundancy of Principia Mathematica". *Philosophy and Phenomenological Research*. **11**: 190–199.

Copi, I. M. (1971). *The Theory of Logical Types*. London: Routledge & Kegan Paul.

Dauben, J. W. (1979). *Georg Cantor: His Mathematics and Philosophy of the Infinite*. Cambridge, MA: Harvard University Press.

Davis, M. (1965). *The Undecidable: Basic Papers on Undecidable Propositions, Unsolvable Problems and Computable Functions*. New York: Raven Press.

Dedekind, R. (1888). *Was sind und sollen die Zahlen?* Braunschweig: Vieweg.

Dedekind, R. (1893). *Was sind und sollen die Zahlen?* Braunschweig: Vieweg. (Segunda edición).

Demopoulos, W., ed. (1995). *Frege's Philosophy of Mathematics*. Cambridge MA: Harvard University Press.

Descartes, R. (1637). *Discours de la méthode pour bien conduire sa raison, et chercher la vérité dans les sciences. Plus la Dioptrique, les Météores et la Géométrie qui sont des essais de cette méthode*. Leyde: Jean Maire.

Detlefsen, M. (1979). "On interpreting Gödel's Second Theorem". En S. G. Shanker 1989, pp. 131–154. (Publicado originalmente, sin el "Postscript", en *Journal of Philosophical Logic*, **8**: 297–313).

Detlefsen, Michael (1986). *Hilbert's Program: An essay on mathematical instrumentalism*. Dordrecht: D. Reidel.

Detlefsen, M. (1990). "On an alleged refutation of Hilbert's Program using Gödel's First Incompleteness Theorem". En M. Detlefsen 1992a, pp. 199-235. (Publicado originalmente en *Journal of Philosophical Logi*c, **18**).

Detlefsen, M., ed. (1992). *Proof and Knowledge in Mathematics*. London: Routledge.

Detlefsen, M., ed. (1992a). *Proof, Logic and Formalization*. London: Routledge.

Dreben, B. (1963). "Corrections to Herbrand". *American Mathematical Society, Notices*. **10**: 285.

Dreben, B. y S. Aanderaa (1964). "Herbrand analyzing functions". *Bulletin of the American Mathematical Society*. **70**: 697–698.

Dreben, B., P. Andrews y S. Aanderaa (1963). "Errors in Herbrand". *American Mathematical Society, Notices*. **10**: 285.

Dreben, B., P. Andrews y S. Aanderaa (1963a). "False lemmas in Herbrand". *Bulletin of the American Mathematical Society*. **69**: 699–706.

Dreben, B. y J. Denton (1966). "A supplement to Herbrand". *Journal of Symbolic Logic*. **31**: 393–398.

Dreben, B. y W. D. Goldfarb (1979). *The Decision Problem. Solvable Classes of Quantificational Formulas*. Reading MA: Addison-Wesley.

du Bois-Reymond, P. (1875). "Über asymptotische Werte, infinitäre Approximationen und infinitäre Auflösungen von Gleichungen". *Mathematische Annalen*. **8**: 363–414.

Etchemendy, John (1990). *The Concept of Logical Consequence*. Cambridge, MA: Harvard University Press.

Euripides (1960). *Bacchae*. Second Edition. Edited with an Introduction and Commentary by E. R. Dodds. Oxford: Clarendon Press.

Felgner, U., ed. (1979). *Mengenlehre*. Darmstadt: Wissenschaftliche Buchgesellschaft.

Fourier, J. (1822). *Théorie analytique de la chaleur.* Paris: Firmin Didot.

Fraenkel, A. A. (1922). "Zu den Grundlagen der Cantor-Zermeloschen Mengenlehre". *Mathematische Annalen*. **86**: 230–237. (Reproducido en Felgner 1979).

Fraenkel, A. A. (1922a). “Der Begriff “definit” und die Unabhängigkeit des Auswahls-axioms”. *Preußische Akademie der Wissenschaften. Physikalisch-mathematische Kl. Sitzungsberichte*. Pp. 253–257. (Traducción inglesa en J. van Heijenoort 1967).

Fraenkel, A. A. (1925). “Untersuchungen über die Grundlagen der Mengenlehre”. *Mathematische Zeitschrift*. **22**: 250–273.

Frege, G. (BSA). *Begriffschrift und andere Aufsätze*. Zweite Auflage mit E. Husserls und H. Scholz’s Anmerkungen. Herausgegeben von I. Angelelli. Darmstadt: Wissenschaftliche Buchgesellschaft, 1971.

Frege, G. (KS). *Kleine Schriften*. Herausgegeben von I. Angelelli. Darmstadt: Wissenschaftliche Buchgesellschaft.

Frege, G. (LS). *Siete escritos sobre lógica y semántica*. Introducción, traducción y selección bibliográfica de Gómez-Lobo, A. Valparaíso: Universidad Católica de Valparaíso, 1972.

Frege, G. (WB). *Wissenschaftlicher Briefwechsel*. Herausgegeben, bearbeitet, eingeleitet und mit Anmerkungen versehen von G. Gabriel, H. Hermes, F. Kambartel, C. Thiel y A. Veraart. Hamburg: Felix Meiner.

Frege, G. (1879). *Begriffschrift, eine der arithmetischen nachgebildete Formelsprache*. Halle a.S.: Louis Nebert.

Frege, G. (1884). *Die Grundlagen der Arithmetik. Eine logisch mathematische Untersuchung über den Begriff der Zahl*. Breslau: Wilhelm Koebner.

Frege, G. (1891). *Funktion und Begriff*. Vortrag, gehalten in der Sitzung vom 9. Januar 1891 der Jenaischen Gesellschaft der Medizin und Naturwissenschaft. Jena: H. Pohle. (Reproducido en G. Frege, KS, pp. 125–142).

Frege, G. (1892). “Über Sinn und Bedeutung”. *Zeitschrift für Philosophie und philosophische Kritik*. **100**: 25–50 (1892). (Reproducido en G. Frege, KS, pp. 143–162).

Frege, G. (1892a). “Über Begriff und Gegenstand”. *Vierteljahrschrift für wissenschaftliche Philosophie*. **16**: 192–205 (1892). (Reproducido en G. Frege, KS, pp. 167–178).

Frege, G. (1893). *Grundgesetze der Arithmetik, begriffschriftlich abgeleitet*. I. Band. Hildesheim: Georg Olms, 1962. (Reimpresión repográfica de la edición original, publicada en Jena en 1893; encuadernada en un volumen con el tomo II).

Frege, G. (1895). “Kritische Beleuchtung einiger Punkte in E. Schröders Vorlesungen über die Algebra der Logik”. *Archiv für systematische Philosophie*. **1**: 433–456. (Reproducido en Frege, KS, pp. 193–210).

Frege, G. (1903). *Grundgesetze der Arithmetik, begriffschriftlich abgeleitet*. II. Band. Hildesheim: Georg Olms, 1962. (Reimpresión repográfica de la edición original, publicada en Jena en 1903, encuadernada en un volumen con el tomo I).

Frege, G. (1903a). "Über die Grundlagen der Geometrie". *Jahresbericht der Deutschen Mathematiker-Vereinigung*. **12**: 319–324. (Reproducido en G. Frege, KS, pp. 262–266).

Frege, G. (1903b). "Über die Grundlagen der Geometrie, II". *Jahresbericht der Deutschen Mathematiker-Vereinigung*. **12**: 368–375. (Reproducido en G. Frege, KS, pp. 267–272).

Frege, G. (1959). *The Foundations of Arithmetic: A Logico-Mathematical Enquiry into the Concept of Number.* Second revised edition. English translation, with facing German text, by J. L. Austin. Oxford: Basil Blackwell.

Freudenthal, H. (1974). "The impact of von Staudt's foundations of geometry". En R. S. Cohen y otros, eds., *For Dirk Struik*. Dordrecht: Reidel. Pp. 189-200.

Geach, P. (1956). "On Frege's way out". En Geach 1972, pp. 235-237. (Publicado originalmente en *Mind*, **65**).

Geach, P. (1972). *Logic Matters*. Oxford: Basil Blackwell.

Gentzen, G. (CP). *The Collected Papers of Gerhard Gentzen*. Translated by M. E. Szabo. Amsterdam: North-Holland.

Gentzen, G. (1934). "Untersuchungen über das logische Schließen". *Mathematische Zeitschrift*. **39**: 176-210, 405-431.

Gentzen, G. (1935). "Die Widerspruchsfreiheit der reinen Zahlentheorie". *Mathematische Annalen*. **112**: 493-565.

Gentzen, G. (1936/37). "Der Unendlichkeitsbegriff in der Mathematik". *Semester-Berichte, Münster in W.* 9. Semester, Winter 1936/37: 65-80. (Traducción inglesa en Gentzen, CP, pp. 223-233).

Gentzen, G. (1938). "Neue Fassung des Widerspruchsfreiheitsbeweises für die reine Zahlentheorie". *Forschungen zur Logik und zur Grundlegung der exakten Wissenschaften*. Neue Folge, **4**: 19-44 (1938).

Gentzen, G. (1943). "Beweisbarkeit und Unbeweisbarkeit von Anfangsfällen der transfiniten Induktion in der reinen Zahlentheorie". *Mathematische Annalen*. **113**: 140-161.

Gergonne, J. D. (1818). "Essai sur la théorie des définitions". *Annales des mathématiques*. **9**: 1–35.

Gillies, D. A. (1982). *Frege, Dedekind, and Peano on the Foundations of Arithmetic*. Assen: Van Gorcum.

Gödel, K. (CW). *Collected Works*. Volume I: Publications 1929–1936; Volume II: Publications 1938–1974; Volume III: Unpublished essays and lectures. Edited by S. Feferman et al. New York: Oxford University Press, 1986–95.

Gödel, K. (OC). *Obras completas*. Segunda edición. Introducción y traducción de J. Mosterín. Madrid: Alianza, 1989.

Gödel, K. (1929). "Über Vollständigkeit des Logikkalküls". En Gödel CW, vol. I, pp. 60–122. (Tesis doctoral, Universidad de Viena).

Gödel, K. (1930) "Die Vollständigkeit der Axiome des logischen Funktionenkalküls". *Monatshefte für Mathematik und Physik*. **37**: 349–360. (Reproducido en Gödel, CW, vol. I, pp. 102–122).

Gödel, K. (1930a). "Einige metamathematische Resultate über Entscheidungs-definitheit und Widerspruchsfreiheit". *Anzeiger derAkademie der Wissenschaften in Wien*. **67**: 214–215 (1930). (Reproducido en Gödel, CW, vol. I, pp. 140–142).

Gödel, K. (1931). "Über formal unentscheidbare Sätze der *Principia Mathematica* und verwandten Systeme". *Monatshefte für Mathematik und Physik*. **7**: 173–198. (Reproducido en Gödel, CW, vol. I, pp. 145-195).

Gödel, K. (1932). "Ein Spezialfall des Entscheidungsproblems der theoretischen Logik". *Ergebnisse eines mathematischen Kolloquiums*. **2**: 27–28 (1932). (Reproducido en Gödel, CW, vol. I, pp. 230–234).

Gödel, K. (1934). "On undecidable propositions of formal mathematical systems". En Gödel CW, vol. I, pp. 346–371. (Notas tomadas en clase por S. C. Kleene y J. B. Rosser, mimeografiadas en 1934 e impresas por primera vez, con correcciones de Gödel, en Davis 1965).

Gödel, K. (1940). *The Consistency of the Axiom of Choice and of the Generalized Continuum Hypothesis with the Axioms of Set Theory.* Princeton: Princeton University Press. (Reproducido en Gödel, CW, vol. II, pp. 33-101).

Gödel, K. (1944). "Russell's mathematical logic". En Gödel CW, vol. II, pp. 119–141. (Publicado originalmente en P. A. Schilpp, ed. *The Philosophy of Bertrand Russell*, Evanston: Northwestern University Press).

Gödel, K. (1958). "Über eine bisher nicht benutzte Erweiterung des finiten Standpunktes". *Dialectica*. **12**: 280-287. (Reproducido en Gödel, CW, vol. II, pp. 240-251).

Goldfarb, W. (1979). "Logic in the twenties: the nature of the quantifier". *Journal of Symbolic Logic*. **44**: 351–368.

Grattan-Guinness, I. (1974). "The rediscovery of the Cantor-Dedekind correspondence". *Jahresbericht der Deutschen Mathematiker-Vereinigung*. **76**: 104–139.

Grelling, K. y L. Nelson (1908). “Bemerkungen zu den Paradoxien von Russell und Burali-Forti”. *Abhandlungen der Fries’schen Schule*. **2**: 301-324.

Hall, A. R. y M. B. Hall, eds. (1962). *Unpublished Scientific Papers of Isaac Newton*. A Selection from the Portsmouth Collection in the University Library, Cambridge. Cambridge: Cambridge University Press.

Hallett, M. (1984). *Cantorian Set Theory and Limitation of Size*. Oxford: Clarendon Press.

Hartogs, F. (1915). “Über das Problem der Wohlordnung”. *Mathematische Annalen*. **76**: 436–443.

Hausdorff, F. (1914). *Grundzüge der Mengenlehre*. Leipzig: Veit.

Henkin, L. (1949). “The completeness of the first-order functional calculus”. *Journal of Symbolic Logic*. **14**: 159-166. (Reproducido en Hintikka 1969, pp. 42-50).

Henkin, L. (1950). “Completeness in the theory of types”. *Journal of Symbolic Logic*. **15**: 81–91 (1950). (Reproducido en Hintikka 1969, pp. 51–63).

Herbrand, J. (EL). *Écrits logiques*. Avec une Préface de J. van Heijenoort, une Notice Biographique par C. Chevalley et A. Lautmann et une Note sur la pensée de Herbrand par C. Chevalley. Paris: Presses Universitaires de France, 1968.

Herbrand, J. (LW). *Logical Writings*. A Translation of the ‘Écrits logiques’, edited by J. van Heijenoort. Edited by W. D. Goldfarb. Cambridge, MA: Harvard University Press, 1971.

Herbrand, J. (1929). “Non-contradiction des axiomes arithmétiques”. *Comptes rendus de l'Académie des Sciences*. **188**: 303-304. (Reproducido en Herbrand, EL, pp. 221–232).

Herbrand, J. (1930). *Recherches sur la théorie de la démonstration*. Varsovia: Prace Towarzystwa Naukowego Warszawskiego. Wydział III, nº 33. (Reproducido en Herbrand, EL, pp. 221–232).

Herbrand, J. (1931). “Sur la non-contradiction de l’arithmétique”. *Journal für die reine und angewandte Mathematik*. **166**: 1-8. (Reproducido en Herbrand, EL, pp. 221–232).

Hermes, H. (1961). *Aufzählbarkeit, Entscheidbarkeit, Berechenbarkeit: Einführung in die Theorie der rekursiven Funktionen*. Berlin: Springer.

Hessenberg, G. (1906). *Grundbegriffe der Mengenlehre*. Abhandlungen der Friesschen Schule, N.R. [I] 4.

Hilbert, D. (GA). *Gesammelte Abhandlungen*. New York: Chelsea, 1965. 3 vols. (Reimpresión de la edición original: Berlin, Springer, 1933–35).

Hilbert, D. (GG). *Grundlagen der Geometrie*. Zehnte Auflage. Mit Supplementen von P. Bernays. Stuttgart: Teubner. (Primera edición, 1899).

Hilbert, D. (GG[7]). *Grundlagen der Geometrie*. Siebente umgearbeitete und vermehrte Auflage. Leipzig: Teubner, 1930.

Hilbert, D. (Hilbertiana). *Hilbertiana. Fünf Aufsätze*. Darmstadt: Wissenschaftliche Buchgesellschaft, 1964.

Hilbert, D. (1899). "Grundlagen der Geometrie". En *Festschrift zur Feier der Enthüllung des Gauss-Weber-Denkmals in Göttingen*. Leipzig: Teubner. Pp. 3–92.

Hilbert, D. (1900). "Über den Zahlbegriff". En Hilbert GG[7], pp. 241-246. (Publicado originalmente en *Jahresbericht der Deutschen Mathematischen Vereinigung*, **8**: 180-194 (1900)).

Hilbert, D. (1900a). "Mathematische Probleme". En D. Hilbert GA, vol. III, . (Publicado originalmente en *Göttinger Nachrichten,* 1900, pp. 253–297; reproducido en *Archiv für Mathematik und Physik*, (3) **1**: 44–63, 213–237 (1901)).

Hilbert, D. (1904). "Über die Grundlagen der Logik und der Arithmetik". En Hilbert GG[7], pp. 247-261. (Publicado originalmente en *Verhandlungen des Dritten Internationalen Mathematiker-Kongresses in Heidelberg vom 8. bis 13. August 1904*, Leipzig: Teubner, 1905, pp. 174-185. Traducción inglesa en van Heijenoort 1967, pp. 129–138).

Hilbert, D. (1922). "Neubegründung der Mathematik (Erste Mitteilung)". *Abhandlungen aus dem mathematischen Seminar der Hamburgischen Universität*. **1**: 155–177. (Reproducido en Hilbert GA, III, 157–177 y en Hilbertiana).

Hilbert, D. (1923). "Die logischen Grundlagen der Mathematik". *Mathematische Annalen*. **88**: 151–165. (Reproducido en Hilbert GA, III, 178–191 y en Hilbertiana).

Hilbert, D. (1926). "Über das Unendliche". *Mathematische Annalen*. **95**: 161-190. (Reproducido en Hilbertiana y en GG[7]; traducción inglesa en van Heijenoort 1967).

Hilbert, D. (1928). "Die Grundlagen der Mathematik". En Hilbert GG[7], pp. 289-312. (Versión ligeramente abreviada del texto publicado originalmente en *Abhandlungen aus dem mathematischen Seminar der Hamburgischen Universität*, **6**: 65-85 (1928)).

Hilbert, D. (1928a). "Probleme der Grundlegung der Mathematik". En Hilbert GG[7], pp. 313-323. (Conferencia leída en el Congreso Internacional de Matemáticos de Bologna el 3 de septiembre de 1928 y publicada en el tomo VI de sus Actas; reproducida en *Mathematische Annalen*, **102**: 1-9 (1930)).

Hilbert, D. (1931). "Zur Grundlegung der elementaren Zahlenlehre". *Mathematische Annalen*. **104**: 485–494.

Hilbert, D. (1931a). “Beweis des Tertium non datur”. *Göttinger Nachrichten (Math.-phys. Klasse)*. Pp. 120–125.

Hilbert, D. y W. Ackermann (1928). *Grundzüge der theoretischen Logik*. Berlin: Springer.

Hilbert, D. y P. Bernays (GM). *Grundlagen der Mathematik*. Zweite Auflage. Berlin: Springer, 1968/1970. 2 vols.

Hilbert, D. y P. Bernays (1934). *Grundlagen der Mathematik*. Band I. Berlin: Springer.

Hilbert, D. y P. Bernays (1939). *Grundlagen der Mathematik*. Band II. Berlin: Springer.

Hintikka, J., ed. (1969). *The Philosophy of Mathematics*. Oxford: Oxford University Press. (Oxford Readings in Philosophy).

Isaacson, D. (1992). “Some considerations on arithmetical truth and the ω-rule”. En M. Detlefsen 1992a, pp. 94-138.

Jaśkowski, S. “On the rules of suppositions in formal logic”. *Studia Logica*. **1**: 5–32 (1934). (Reproducido en McCall 1967, pp. 231-258).

Jeffrey, R. (1981). *Formal Logic: Its Scope and Limits*. Second edition. New York: McGraw-Hill.

Johnson, D. M. (1979/81). “The problem of the invariance of dimension in the growth of modern topology”. *Archive for History of Exact Sciences*. **20**: 97–188; **25**: 85–267.

Kalmár, L. (1955). “Über ein Problem, betreffend die Definition des Begriffes der allgemein-rekursiven Funktionen”. *Zeitschrift für mathematische Logik und Grundlagenforschung*. **1**: 93-96.

Kleene, S. C. (1936). “General recursive functions of natural numbers”. *Mathematische Annalen*. **112**: 727-742. (Reproducido en Davis 1965, pp. 237-252, con importantes rectificaciones en la p. 253. Mis citas remiten a esta edición.).

Kleene, S. C. (1938). “On notation for ordinal numbers”. *Journal of Symbolic Logic*. **3**: 150–155.

Kleene, S. C. (1943). “Recursive predicates and quantifiers”. *American Mathematical Society Transactions*. **53**: 41-73 (1943). (Reproducido en Davis 1965, pp. 255-287. Mis citas remiten a esta edición.).

Kleene, S. C. (1952). *Introduction to Metamathematics*. Amsterdam: North-Holland.

Kleene, S. C. (1967). *Mathematical Logic*. New York: Wiley.

König, D. (1926). “Sur les correspondances multivoques des ensembles”. *Fundamenta Mathematicae*. **8**: 114-134.

König, J. (1905). “Zum Kontinuum-Problem”. *Mathematische Annalen*. **60**: 177–180, 462.

König, J. (1905a). "Über die Grundlagen der Mengenlehre und das Kontinuumsproblem". *Mathematische Annalen*. **61**: 156-160. (Traducción inglesa en van Heijenoort 1967, pp. 145-149).

König, J. (1914). *Neue Grundlagen der Logik, Arithmetik und Mengenlehre*. Leipzig.

Kreisel, G. (1976). "What have we learnt from Hilbert's Second Problem". En F. E. Browder 1976, pp. 93–130.

Kronecker, L. (1887). "Über den Zahlbegriff". *Journal für die reine und angewandte Mathematik*. **101**: 337–355.

Largeault, J. (1992). *L'intuitionisme*. Paris: Presses Universitaires de France.

Leibniz, G. W. (EF). *Escritos filosóficos*. Edición de E. de Olaso. Buenos Aires: Charcas, 1982.

Leibniz, G. W. (GP). *Die philosophischen Schriften*. Herausgegeben von C. J. Gerhardt. Hildesheim: Olms, 1965. 7 vols.

Leibniz, G. W. (LP). *Logical Papers*. A Selection Translated and Edited with an Introduction by G. H. R. Parkinson. Oxford: Clarendon Press.

Levy, A. (1979). *Basic Set Theory.* Berlin: Springer.

Lewis, C. I. (1918). *A Survey of Symbolic Logic*. Berkeley: University of California Press.

Lewis, D. (1991). *Parts of Classes*. With an appendix by J. P. Burgess, A. P. Hazen and D. Lewis. Oxford: Basil Blackwell.

Lorenzen, P. (1965). *Differential und Integral*. Frankfurt a. M.

Löwenheim, L. (1915). "Über Möglichkeiten im Relativkalkül". *Mathematische Annalen*. **76**: 447-470.

Maddy, P. (1990). *Realism in Mathematics*. Oxford: Clarendon Press.

Mancosu, P. (1998). *From Brouwer to Hilbert: The Debate on the Foundations of Mathematics in the 1920s*. New York: Oxford University Press.

Manin, Y. I. (1977). *A Course in Mathematical Logic*. Translated from the Russian by N. Koblitz. New York: Springer.

Mannheim, J. H. (1964). *The Genesis of Point Set Topology.* Oxford: Pergamon.

Mates, B. (1970). *Elementary Logic*. New York: Oxford University Press.

Matijasévič, Y. V. (1970). "Diofantovost pereçislimyh množestv". *Doklady Akad. Nauk SSSR*. **191**: 279-282.

Matijasévič, Y. V. (1970a). "Enumerable sets are Diophantine". *Soviet Math. Doklady.* **11**: 354-357. (Versión inglesa corregida de Matijasévič 1970).

McCall, S., ed. (1967). *Polish Logic 1920–1939*. Papers by Adjukiewicz, Chwistek, Jaśkowski, Jordan, Leśniewski, Łukasiewicz, Slupecki, Sobociński, and Wajsberg. With an Introduction by T. Kotarbinski. Translated by B. Gruchman et al. Oxford: Clarendon Press.

Mendelson, E. (1958). "The Axiom of Fundierung and the Axiom of Choice". *Archiv für mathematische Logik und Grundlagenforschung*. **4**: 65–70.

Mendelson, E. (1990) "Second thoughts about Church's Thesis and mathematical proofs". *Journal of Philosophy*. **87**: 225–233.

Meschkowski, H. (1967). *Probleme des Unendlichen: Werk und Leben Georg Cantors*. Braunschweig: Vieweg.

Mirimanoff, D. (1917). "Les antinomies de Russell et de Burali-Forti et le problème fondamentale de la théorie des ensembles". *L'Enseignement Mathématique*. **19**: 37–52.

Mirimanoff, D. (1917a). "Remarques sur la théorie des ensembles et les antinomies cantoriennes – I". *L'Enseignement Mathématique*. **19**: 209–217.

Mirimanoff, D. (1920). "Remarques sur la théorie des ensembles et les antinomies cantoriennes – II". *L'Enseignement Mathématique*. **21**: 29–52.

Mittelstraß, J., ed. (EPW). *Enzyklopädie Philosophie und Wissenschaftstheorie*. Mannheim: Bibliographisches Institut / Stuttgart: J.B. Metzler. 1980–1996. 4 vols.

Moore, E. H. y H. L. Smith (1922). "A general theory of limits". *American Journal of Mathematics*. **44**: 102-121.

Moore, G. H. (1982). *Zermelo's Axiom of Choice: Its origins, development, and influence*. New York: Springer.

Nagel, E. (1939). "The formation of modern conceptions of formal logic in the development of geometry". *Osiris*. **7**: 142-224.

Nelson, E. (1986). *Predicative Arithmetic*. Princeton: Princeton University Press.

Nelson, E. (1993). "Taking formalism seriously". *The Mathematical Intelligencer*. **15** 3: 8-11.

Neumann, J. von (CW). *Collected Works*. Volume I. New York: Pergamon, 1961.

Neumann, J. von (1922/23). "Zur Einführung der transfiniten Zahlen". *Acta Litterarum ac Scientiarum Regiae Universitatis Hungaricae, Sectio Sc. Math.* : 199–208. (Reproducido en von Neumann, CW, vol. I, pp. 24–33 y en Felgner 1979, pp. 92–101; traducción inglesa en Heijenoort 1967, pp. 346–354).

Neumann, J. von (1925). "Eine Axiomatisierung der Mengenlehre". *Journal für die reine und angewandte Mathematik*. **154**: 219-240 (1925). (Reproducido en von Neumann, CW, vol. I, pp. 35–56).

Neumann, J. von (1927a). “Mathematische Begründung der Quantenmechanik”. *Göttinger Nachrichten*. **1** Nr. 9: 1–57.

Neumann, J. von (1927). “Zur Hilbertschen Beweistheorie”. *Mathematische Zeitschrift*. **26**: 1–46 (1927). (Reproducido en von Neumann, CW, vol. I, pp. 256–300).

Neumann, J. von (1928). “Die Axiomatisierung der Mengenlehre”. *Mathematische Zeitschrift*. **27**: 669–752. (Reproducido en von Neumann, CW, vol. I, pp. 339–422).

Neumann, J. von (1928a). “Über die Definition durch transfinite Induktion und verwandte Fragen der allgemeinen Mengenlehre”. *Mathematische Annalen*. **99**: 373-391. (Reproducido en von Neumann, CW, vol. I, pp. 320–338).

Neumann, J. von (1928b). “Zur Theorie der Gesellschaftsspiele”. *Mathematische Annalen*. **100**: 295–320.

Neumann, J. von (1931). “The formalist foundations of mathematics”. En P. Benacerraf y H. Putnam 1983, pp. 61-65. (Traducido al inglés por Erna Putnam y G. J. Massey del original alemán publicado en *Erkenntnis* en 1931).

Neumann, J. von (1932). *Mathematische Grundlagen der Quantenmechanik*. Berlin: Springer.

Neumann, J. von y O. Morgenstern (1944). *Theory of Games and Economic Behavior.* Princeton: Princeton University Press.

Pasch, M. (1882). *Vorlesungen über neuere Geometrie*. Leipzig: Teubner.

Peano, G. (1889). *Arithmetices principia nova methodo exposita*. Torino: Bocca.

Peano, G. (1895-1908). *Formulaire des Mathématiques*. Torino: Bocca. 5 vols.

Peano, G. (1894). *Notations de Logique Mathématique: Introduction au Formulaire de Mathématique publiée par Rivista di Matematica*. Turin: Guadagni.

Peano, G. (1906). “Super theorema de Cantor-Bernstein”. *Rendiconti del Circolo matematico di Palermo*. **21**: 360.

Peirce, C. S. (CP). *Collected Papers*. Edited by C. Hartshorne, P. Weiss y A. W. Burks. Cambridge MA: The Belknap Press of Harvard University Press. 8 vols.

Peirce, C. S. (1880). “On the algebra of logic”. En Peirce CP, vol. 3, pp. 104–157. (Publicado originalmente en *American Journal of Mathematics*, **3**: 15–97).

Peirce, C. S. (1882). “Brief description of the algebra of relatives”. En Peirce CP, vol. 3, pp. 180–186. (Publicado póstumamente aquí por primera vez).

Peirce, C. S. (1883). “A theory of probable inference. Note B. The logic of relatives”. En Peirce CP, vol. 3, pp. 195–209. (Publicado originalmente en *Studies in Logic by Members of the Johns Hopkins University*, Boston, pp. 187–203).

Peirce, C. S. (1870). “Description of a notation for the logic of relatives, resulting from an amplification of the conceptions of Boole’s calculus of logic”. En Peirce CP, vol. 3, pp. 27–98. (Publicado originalmente en *Memoirs of the American Academy of Arts and Sciences*, **9**: 317–78).

Peirce, C. S. (1901). “Insolubilia”. En Peirce CP, vol. 2, pp. 370-371. (Publicado originalmente en Baldwin 1901-1905, vol. I, p. 554).

Péter, R. (1934). “Über den Zusammenhang der verschiedener Begriffe der rekursiven Funktion”. *Mathematische Annalen*. **110**: 612-632.

Poincaré, H. (CM). *Ciencia y método*. Traducción castellana por M. García Miranda y L. Alonso. Madrid: Espasa-Calpe, 1963. (Traducción de Poincaré 1908).

Poincaré, H. (DP). *Dernières Pensées*. Paris: Flammarion, 1963. (Primera edición, 1913).

Poincaré, H. (SH). *La science et l’hypothése*. Paris: Flammarion, 1968. (Primera edición, 1902).

Poincaré, H. (1894). “Sur la nature du raisonnement mathématique”. *Revue de métaphysique et de morale*. **2**: 371–384.

Poincaré, H. (1905/1906). “Les mathématiques et la logique”. *Revue de Métaphysique et de Morale*. **13**: 815–835; **14**: 17–34, 294–317.

Poincaré, H. (1908). *Science et méthode*. Paris: Flammarion.

Poincaré, H. (1909). “La logique de l’infini”. *Revue de Métaphysique et de Morale*. **17**: 461–482.

Poncelet, J. V. (1822). *Traité des propriétés projectives des figures*. Paris: Bachelier.

Pollack, S. (1982). “The development of computer science”. En Pollack SCS, pp. 1–51.

Pollack, S. V., ed. (SCS). *Studies in Computer Science*. Washington: Mathematical Association of America, 1982.

Pollard, S. (1990). *Philosophical Introduction to Set Theory*. Notre Dame: University of Notre Dame Press.

Poncelet, J. V. (1822). *Traité des propriétés projectives des figures*. Paris: Bachelier.

Post, E. L. (1921). “Introduction to a general theory of elementary propositions”. En van Heijenoort 1967, pp. 264–283. (Publicado originalmente en *Bulletin of the American Mathematical Society*, **43**: 163–185 (1921)).

Post, E. L. (1936). “Finate combinatory processes: Formulation I”. *Journal of Symbolic Logic*. **1**: 103-105 (1936). (Reproducido en Davis 1965, pp. 289-291. Mis referencias remiten a esta edición).

Post, E. L. (1944). “Recursively enumerable sets of positive integers and their decision problems”. *Bulletin of the American Mathematical Society*. **50**: 284-316 (1944). (Reproducido en Davis 1965, pp. 305-337. Mis citas remiten a esta edición).

Post, E. L. (1965). “Absolutely unsolvable problems and relatively undecidable propositions: Account of an anticipation”. En Davis 1965, pp. 340–433.

Quine, W. V. O. (1941). “Whitehead and the rise of modern logic”. En Quine 1966, pp. 3–36. (Publicado originalmente en P. A. Schilpp, ed. *The Philosophy of Alfred North Whitehead*, Evanston: Northwestern University Press).

Quine, W. V. O. (1955). “On Frege’s way out”. En Quine 1966, pp. 146-158. (Publicado originalmente en *Mind*, **64**: 145-159).

Quine, W. V. O. (1966). *Selected Logic Papers*. New York: Random House.

Ramsey, F. P. (1925). “The foundations of mathematics”. En Ramsey 1931, pp. 1–61. (Publicado originalmente en *Proceedings of the London Mathematical Society* (2) **25**: 338–384).

Ramsey, F. P. (1931). *The Foundations of Mathematics and other Logical Essays*. London: Routledge & Kegan Paul.

Rang, B. y W. Thomas (1980). “Zermelo’s discovery of the ‘Russell paradox’”. *Historia Mathematica*. **8**: 15–22.

Resnik, M. D. (1980). *Frege and the Philosophy of Mathematics*. Ithaca: Cornell University Press.

Richard, J. (1905). “Les principes des mathématiques et le problème des ensembles”. *Revue générale des sciences pures et appliquées*. **16**: 541 (1905). (Traducción inglesa en van Heijenoort 1967, pp. 142–144).

Ritter, J. y K. Gründer, eds. (HWB). *Historisches Wörterbuch der Philosophie*. Darmstadt: Wissenschaftliche Buchgesellschaft, 1971– .

Robinson, A. (1961). “Non-standard analysis”. *K. Nederlandse Akademie van Wetenschappen. Proceedings*. **64**: 432-440.

Robinson, A. (1966). *Non-standard Analysis*. Amsterdam: North-Holland Publishing Co.

Rodríguez-Consuegra, F. A. (1989). “Russell’s Theory of Types, 1901-1910: Its complex origins in the unpublished manuscripts”. *History and Philosophy of Logic*. **10**: 131-164.

Rodríguez-Consuegra, F. A. (1991). *The Mathematical Philosophy of Bertrand Russell: Origins and Development*. Basel: Birkhäuser.

Rogers Jr., H. (1967). *Theory of Recursive Functions and Effective Computability.* New York: McGraw-Hill.

Rosser, J. B. (1936). “Extensions of some theorems of Gödel and Church”. *Journal of Symbolic Logic*. **1**: 87-91 (1936). (Reproducido en Davis 1965, pp. 231-235. Mis citas remiten a esta edición).

Russell, B. (EA). *Essays in Analysis*. Edited by D. Lackey. London: Allen & Unwin, 1973.

Russell, B. (LK). *Logic and Knowledge*. *Essays 1901-1950*. Edited by R. C. Marsh. London: Allen and Unwin.

Russell, B. (1903). *The Principles of Mathematics*. Volume I. Cambridge: Cambridge University Press.

Russell, B. (1906). “On some difficulties in the theory of transfinite numbers and order types”. *London Mathematical Society Proceedings*. **4**: 29–53. (Reproducido en Russell EA, pp. 135-164).

Russell, B. (1906a). “On the substitutional theory of classes and relations”. En Russell EA, pp. 165-189. (Publicado postumamente en 1973).

Russell, B. (1906b). “Les paradoxes de la logique”. *Revue de métaphysique et de morale*. **14**: 627–650 (1906). (El original inglés es Russell 1906c).

Russell, B. (1906c). “On ‘insolubilia’ and their solution by symbolic logic”. En Russell EA, pp. 190–214. (Este es el original inglés de Russell 1906b, publicado postumamente en 1973).

Russell, B. (1908). “Mathematical logic as based on the theory of types”. *American Journal of Mathematics*. **30**: 222-262.

Russell, B. (1910). “La théorie des types logiques”. *Revue de métaphysique et de morale*. **18**: 263-301. (El original inglés, titulado “The theory of logical types”, se publicó postumamente en Russell EA, pp. 215-252).

Russell, B. (1919). *Introduction to Mathematical Philosophy*. London: Allen & Unwin.

Russell, B. (1945). *Introducción a la filosofía matemática*. Traducción de J.B. Molinari. Anotada y revisada por F. D. Jaime. Buenos Aires: Losada.

Russell, B. (1937). *The Principles of Mathematics*. Second edition. London: Allen & Unwin.

Schilpp, P. A. (1944). *The Philosophy of Bertrand Russell*. Evanston IL: Northwestern University Press. (Library of Living Philosophers).

Scholz, H. (1930). “Die Axiomatik der Alten”. *Blätter für deutsche Philosophie*. **4**: 259–278.

Schröder, E. (VAL). *Vorlesungen über die Algebra der Logik (Exakte Logik)*. New York: Chelsea, 1966. 3 vols. (Reimpresión corregida de la edición original, publicada en Leipzig, 1890–1905).

Schröder, E. (1877). *Der Operationskreis des Logikkalkuls*. Leipzig: Teubner. (Reimpresión inalterada: Darmstadt, Wissenschaftliche Buchgesellschaft, 1966).

Schütte, K. (1951). "Beweistheoretische Erfassung der unendlichen Induktion in der Zahlentheorie". *Mathematische Annalen*. **122**: 369–389.

Schütte, K. (1960). *Beweistheorie*. Berlin: Springer.

Shanker, S. G., ed. (1989). *Gödel's Theorem in Focus*. London: Routledge.

Shoenfield, J. R. (1967). *Mathematical Logic*. Reading, MA: Addison-Wesley.

Skolem, T. (SWL). *Selected Works in Logic*. Edited by J. E. Fenstad. Oslo: Universitetsforlaget, 1970.

Skolem, T. (1920). "Logisch-kombinatorische Untersuchungen über die Erfüllbarkeit und Beweisbarkeit mathematischen Sätze nebst einem Theorem über dichte Mengen". En Skolem SWL, pp. 103–136. (Publicado originalmente en Videnskabsakademiet i Kristiania, *Skrifter*, I, Nº 4, pp. 1-36).

Skolem, T. (1922). "Einige Bemerkungen zur axiomatischen Begründung der Mengenlehre". En Skolem SWL, pp. 137–152. (Publicado originalmente en Proceedings of the 5th Scandinavian Math. Congress, Helsinki, 1922, pp. 217–232).

Skolem, T. (1923). "Begründung der elementären Arithmetik durch die rekurrierende Denkweise ohne Anwendung scheinbarer Veränderlichen mit unendlichem Ausdehnungsbereich". En Skolem SWL, pp. 153–188. (Publicado originalmente en Videnskabsakademiet i Kristiania, *Skrifter*, I, Nº 6).

Skolem, T. (1930). "Einige Bemerkungen zu der Abhandlung von E. Zermelo: "Über die Definitheit in der Axiomatik".". En Skolem SWL, pp. 275–279. (Publicado originalmente en *Fundamenta Mathematicae*, **15**: 337–341 (1930)).

Smorynski, C. (1977). "The incompleteness theorems". En J. Barwise, ed. *Handbook of mathematical logic*. Amsterdam: North-Holland Publishing Co., 1977. Pp. 821-865.

Smorynski, C. (1985). *Self-Reference and Modal Logic*. New York: Springer Verlag.

Smullyan, R. M. (1992). *Gödel's Incompleteness Theorems*. New York: Oxford University Press. xiii + 139 pp.

Soare, R. I. (1987). *Recursively Enumerable Degrees: A Study of Computable Functions and Computably Generated Sets*. Berlin: Springer.

Sobociński, B. (1949). "L'analyse de l'antinomie russellienne para Leśniewski". *Methodos*. **1**: 94-107, 220-228, 308-316.

Takeuti, G. (1975). *Proof Theory*. Amsterdam: North-Holland.

Tiles, M. (1991). *Mathematics and the Image of Reason*. London: Routledge.

Torretti, R. (1978). *Philosophy of Geometry from Riemann to Poincaré*. Dordrecht: D. Reidel.

Torretti, R. (1993). "El método axiomático". En *La ciencia: estructura y desarrollo*. Edición de C. U. Moulines. Madrid: Trotta. Pp. 89-110. (*Enciclopedia Iberoamericana de Filosofía*, vol. 4).

Turing, A. M. (1936). "On computable numbers, with an application to the *Entscheidungsproblem*". *London Mathematical Society Proceedings*. **42**: 230-265.. (Reproducido en Davis 1965, pp. 116-151).

Turing, A. M. (1937). "On computable numbers, with an application to the *Entscheidungsproblem*: A correction". *London Mathematical Society Proceedings*. **43**: 544-546 (1937). (Reproducido en Davis 1965, pp. 152-154).

van Dalen, D. (1990). "The War of the Frogs and the Mice, or the Crisis of the Mathematische Annalen". *Mathematical Intelligencer*. **12**, 4: 17-31.

van Heijenoort, J. (1967). *From Frege to Gödel: A Source Book in Mathematical Logic, 1879-1931*. Cambridge MA: Harvard University Press.

van Stigt, W. P. (1979). "The rejected parts of Brouwer's dissertation". *Historia Mathematica*. **6**: 385-404.

Wang, Hao (1987). *Reflections on Kurt Gödel*. Cambridge MA: The MIT Press.

Webb, J. C. (1980). *Mechanism, Mentalism, and Metamathematics: An Essay on Finitism*. Dordrecht: D. Reidel.

Weil, A. (1937). *Sur les espaces à structure uniforme et sur la topologie générale*. Paris: Hermann.

Weyl, H. (GA). *Gesammelte Abhandlungen*. Herausgegeben von K. Chandrasekharan. Berlin: Springer, 1968. 4 vols.

Weyl, H. (1910). "Über die Definitionen der mathematischen Grundbegriffe". En Weyl GA, vol. I, pp. 298–304. (Publicado originalmente en *Mathematisch-naturwissenschaftliche Blätter*, **7**: 93–95, 109–113 (1910)).

Weyl, H. (1918). *Das Kontinuum. Kritische Untersuchungen über die Grundlagen der Analysis*. New York: Chelsea, 1973. (Reimpresión de la obra aparecida originalmente en Leipzig en 1918).

Weyl, H. (1921). "Über die neue Grundlagenkrise der Mathematik". *Mathematische Zeitschrift*. **10**: 39–79.

Whitehead, A. N. y B. Russell (1910-13). *Principia Mathematica*. Cambridge: Cambridge University Press. 3 vols.

Wittgenstein, L. (PB). *Philosophische Bemerkungen*. Aus dem Nachlaß herausgegeben von R. Rhees. Frankfurt/M: Suhrkamp, 1964.

Wittgenstein, L. (1922). *Tractatus Logico-Philosophicus*. With a facing English translation by C. K. Ogden. London: Routledge & Kegan Paul.

Wittgenstein, L. (1956). *Bemerkungen über die Grundlagen der Mathematik*. Herausgegeben und bearbeitet von G. H. von Wright, R. Rhees and G. E. M. Anscombe. Oxford: Basil Blackwell. (With a facing English translation by G. E. M. Anscombe).

Wittgenstein, L. (1974). *Letters to Russell, Keynes and Moore*. Edited with an introduction by G. H. von Wright, Ithaca: Cornell University Press.

Wright, C. (1983). *Frege's Conception of Numbers as Objects*. Aberdeen: Aberdeen University Press.

Zermelo, E. (1904). "Beweis, daß jede Menge wohlgeordnet werden kann". *Mathematische Annalen*. **59**: 514-516.

Zermelo, E. (1908). "Neuer Beweis für die Möglichkeit einer Wohlordnung". *Mathematische Annalen*. **65**: 514-535.

Zermelo, E. (1908a). "Untersuchungen über die Grundlagen der Mengenlehre I". *Mathematische Annalen*. **65**: 261-281.

Zermelo, E. (1929). "Über den Begriff der Definitheit in der Axiomatik". *Fundamenta Mathematicae*. **14**: 339–344.

Zermelo, E. (1930). "Über Grenzzahlen und Mengenbereiche". *Fundamenta Mathematicae*. **16**: 29–47.

# ABREVIATURAS Y SÍMBOLOS

## *Símbolos asociados a determinados autores*

# ÍNDICE DE PERSONAS Y CONCEPTOS

La numeración de las páginas indicadas corresponde a la edición impresa en 1998 y en algunos casos se aparta —aunque solo mínimamente— del presente pdf.

El índice no cubre la lista de obras citadas. En algunas partidas correspondientes a conceptos figuran números de página seguidos de la letra n. Esta letra indica que en una nota al pie de la página mencionada hay una referencia importante al concepto respectivo. La letra n se omite cuando el concepto pertinente se menciona en el texto principal de la página, aunque además se refiera a él una nota al pie de la misma. En las partidas correspondientes a nombres propios normalmente no se indica la región de la página en que aparece el nombre respectivo.

Imitando al *Diccionario* de María Moliner, el orden alfabético seguido aquí prescinde de la curiosa convención castellana, que trata la combinación ‘ch’ como una sola letra, aunque a todas luces son dos; por ende, ‘ch’ va entre ‘ce’ y ‘ci’.